



# Progressive Parameterizations

LIGANG LIU, University of Science and Technology of China, China

CHUNYANG YE, University of Science and Technology of China, China

RUIQI NI, University of Science and Technology of China, China

XIAO-MING FU\*, University of Science and Technology of China, China

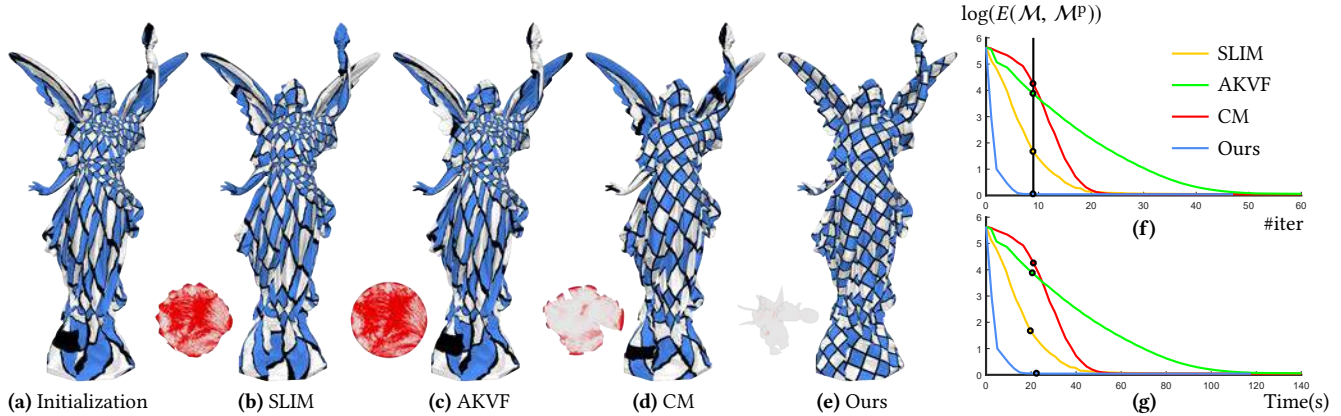


Fig. 1. Parameterizing a disk topology mesh (1792000 triangles) with low isometric distortion and no foldovers by optimizing the symmetric Dirichlet energy ( $E(\mathcal{M}, \mathcal{M}^P)$  in Problem (5)). Starting from the same bijective initialization [Tutte 1963] (a), the energy is measured during the optimization process. The figure (b - e) shows a snapshot of the state each method achieved at the 9<sup>th</sup> iteration marked on the graph (f). The corresponding time in seconds is marked on the graph (g). The color of the triangles from the parameterized meshes encodes the symmetric Dirichlet energy metric, with white being optimal. Our approach achieves much better efficiency than the competitors including SLIM [Rabinovich et al. 2017], AKVF [Claici et al. 2017], and CM [Shtengel et al. 2017]. In order to achieve our result in (e), SLIM, AKVF and CM require 144, 81, and 17 more iterations and 314.60, 184.45, and 38.65 more seconds, respectively.

We propose a novel approach, called *Progressive Parameterizations*, to compute foldover-free parameterizations with low isometric distortion on disk topology meshes. Instead of using the input mesh as a reference to define the objective function, we introduce a progressive reference that contains bounded distortion to the parameterized mesh and is as close as possible to the input mesh. After optimizing the bounded distortion energy between the progressive reference and the parameterized mesh, the parameterized mesh easily approaches the progressive reference, thereby also coming close to the input. By iteratively generating the progressive reference and optimizing the bounded distortion energy to update the parameterized mesh, our algorithm achieves high-quality parameterizations with strong practical reliability and high efficiency. We have demonstrated that our algorithm succeeds on a massive test data set containing over 20712 complex disk topology meshes. Compared to the state-of-the-art methods, our method has achieved higher computational efficiency and practical reliability.

\*The corresponding author

Authors' addresses: Ligang Liu, University of Science and Technology of China, China, lgliu@ustc.edu.cn; Chunyang Ye, University of Science and Technology of China, China, yechyang@mail.ustc.edu.cn; Ruiqi Ni, University of Science and Technology of China, China, nrq@mail.ustc.edu.cn; Xiao-Ming Fu, University of Science and Technology of China, China, fuxm@ustc.edu.cn.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART41 \$15.00

<https://doi.org/10.1145/3197517.3201331>

CCS Concepts: • **Computing methodologies** → **Shape modeling**;

Additional Key Words and Phrases: parameterizations, isometric mapping, low distortion, foldover-free, flip-free

## ACM Reference Format:

Ligang Liu, Chunyang Ye, Ruiqi Ni, and Xiao-Ming Fu. 2018. Progressive Parameterizations. *ACM Trans. Graph.* 37, 4, Article 41 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201331>

## 1 INTRODUCTION

Surface parameterization is a fundamental problem in computer graphics and geometric processing; thus, it has been widely used in many applications, such as texture mapping, remeshing, inter-surface mapping, and shape analysis (cf. extensive surveys [Floater and Hormann 2005; Hormann et al. 2007; Sheffer et al. 2006]).

These tasks rely on the computation of a *foldover-free, low-distortion* parameterization. Linear methods, such as Tutte's method and its variants [Aigerman and Lipman 2015; Floater 2003; Tutte 1963], provide injectivity-guaranteed parameterization, but usually exhibit extremely large distortion for complex inputs. Nonlinear methods formulate parameterization as an optimization problem by minimizing an energy functions with some constraints that preserve the orientations of triangles. Generally, the objective function includes a low-distortion term that is large when the triangle is highly distorted. It may also include a foldover-preventing term that goes to infinity when a triangle flips or degenerates. These objective functions are highly non-convex and non-linear; thus, they are numerically difficult to optimize, particularly for large-scale inputs.



# Anderson Acceleration for Geometry Optimization and Physics Simulation

YUE PENG, University of Science and Technology of China  
 BAILIN DENG, Cardiff University  
 JUYONG ZHANG\*, University of Science and Technology of China  
 FANYU GENG, University of Science and Technology of China  
 WENJIE QIN, University of Science and Technology of China  
 LIGANG LIU, University of Science and Technology of China

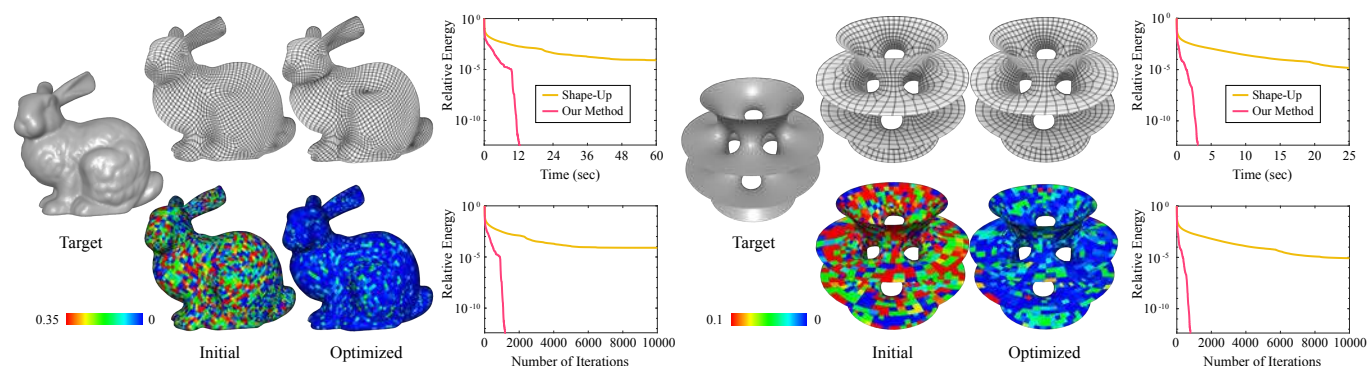


Fig. 1. Compared with the local-global solver from [Bouaziz et al. 2012], our method significantly reduces the computational time and iterations required for an accurate solution to planar quad mesh optimization, as shown in the log-scale relative energy error graphs. The color coding shows the planarity error for each face, computed as the max distance from its vertices and their best fitting plane, normalized by the average edge length of the mesh.

Many computer graphics problems require computing geometric shapes subject to certain constraints. This often results in non-linear and non-convex optimization problems with globally coupled variables, which pose great challenge for interactive applications. Local-global solvers developed in recent years can quickly compute an approximate solution to such problems, making them an attractive choice for applications that prioritize efficiency over accuracy. However, these solvers suffer from lower convergence rate, and may take a long time to compute an accurate result. In this paper, we propose a simple and effective technique to accelerate the convergence of such solvers. By treating each local-global step as a fixed-point iteration, we apply Anderson acceleration, a well-established technique for fixed-point solvers, to speed up the convergence of a local-global solver. To address the stability issue of classical Anderson acceleration, we propose a simple strategy to guarantee the decrease of target energy and ensure its global

convergence. In addition, we analyze the connection between Anderson acceleration and quasi-Newton methods, and show that the canonical choice of its mixing parameter is suitable for accelerating local-global solvers. Moreover, our technique is effective beyond classical local-global solvers, and can be applied to iterative methods with a common structure. We evaluate the performance of our technique on a variety of geometry optimization and physics simulation problems. Our approach significantly reduces the number of iterations required to compute an accurate result, with only a slight increase of computational cost per iteration. Its simplicity and effectiveness makes it a promising tool for accelerating existing algorithms as well as designing efficient new algorithms.

CCS Concepts: • **Computing methodologies** → **Computer graphics; Animation**; • **Theory of computation** → *Nonconvex optimization*;

Additional Key Words and Phrases: Fixed-point iterations, numerical optimization, parallel computing, projective dynamics, geometry processing

## ACM Reference Format:

Yue Peng, Bailin Deng, Juyong Zhang, Fanyu Geng, Wenjie Qin, and Ligang Liu. 2018. Anderson Acceleration for Geometry Optimization and Physics Simulation. *ACM Trans. Graph.* 37, 4, Article 42 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201290>

## 1 INTRODUCTION

Many computer graphics problems require computing geometric shapes whose elements are subject to certain constraints. In geometric modeling, such constraints can be related to aesthetics, performance, or fabrication requirements of the shape. For example, for freeform architectural design represented as quadrilateral meshes,

\*Corresponding author (juyong@ustc.edu.cn).

Authors' addresses: {Yue Peng, Juyong Zhang, Fanyu Geng, Wenjie Qin, Ligang Liu}, University of Science and Technology of China, 96 Jinzhai Road, Hefei, Anhui, 230026, China; {echoyue@mail.ustc.edu.cn, juyong@ustc.edu.cn, gfy29110@mail.ustc.edu.cn, darkqin@mail.ustc.edu.cn, lgliu@ustc.edu.cn}; Bailin Deng, Cardiff University, 5 The Parade, Cardiff CF24 3AA, Wales, United Kingdom, DengB3@cardiff.ac.uk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART42 \$15.00

<https://doi.org/10.1145/3197517.3201290>

# Fast Winding Numbers for Soups and Clouds

GAVIN BARILL, University of Toronto, Canada

NEIL G. DICKSON, Side Effects Software Inc., Canada

RYAN SCHMIDT, Gradientspace, Canada

DAVID I.W. LEVIN\* and ALEC JACOBSON\*, University of Toronto, Canada

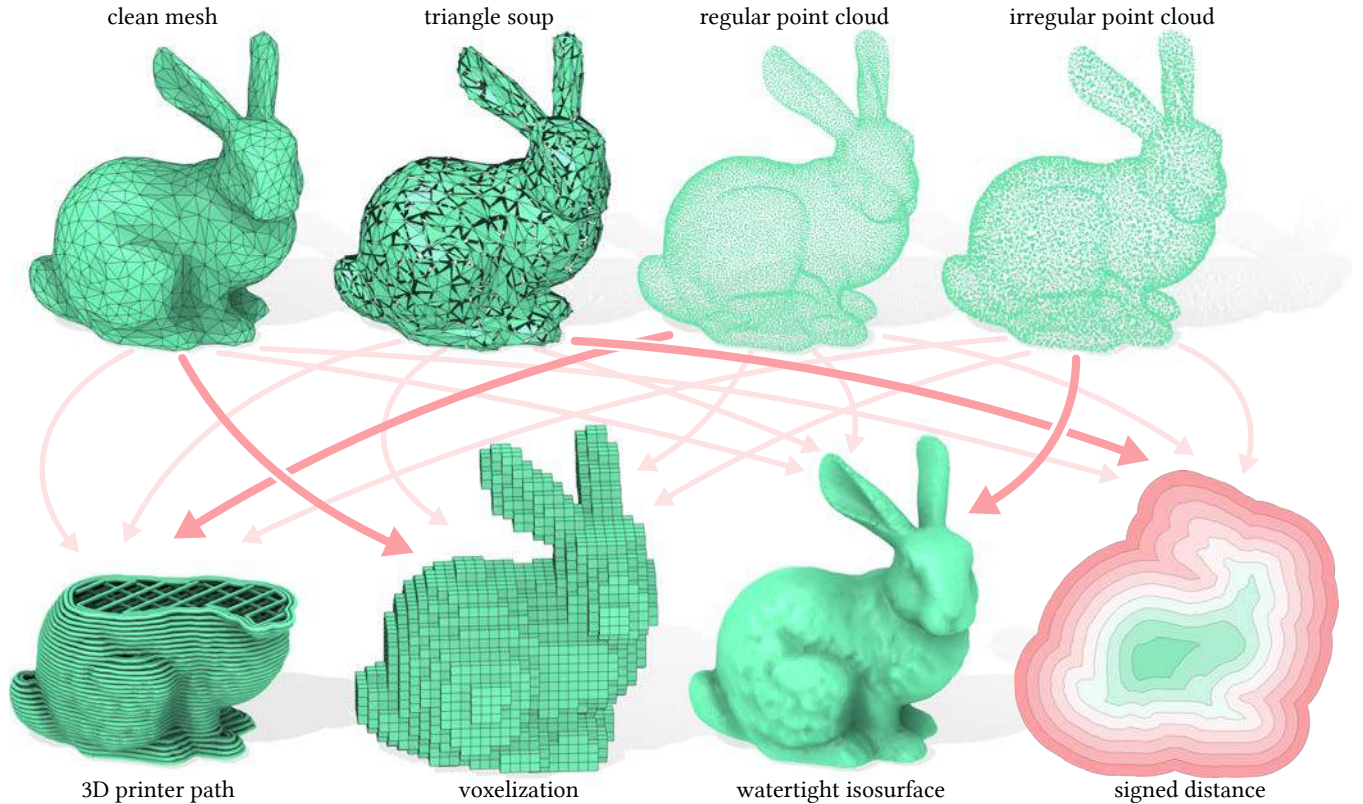


Fig. 1. In this paper, we further generalize the winding number to point clouds and propose a hierarchical algorithm for fast evaluation (up to 1000× speedup). This enables efficient answers to inside-outside queries for a wider class of shape representations (top) during a variety of tasks (bottom).

Inside-outside determination is a basic building block for higher-level geometry processing operations. Generalized winding numbers provide a robust answer for triangle meshes, regardless of defects such as self-intersections, holes or degeneracies. In this paper, we further generalize the winding number to point clouds. Previous methods for evaluating the winding number are

\*Joint last authors

Authors' addresses: Gavin Barill, David I.W. Levin, Alec Jacobson, UofT, 40 St George St, Toronto, M5S 2E4, Canada, gavin.barill@mail.utoronto.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART43 \$15.00

<https://doi.org/10.1145/3197517.3201337>

slow for completely disconnected surfaces, such as triangle soups or—in the extreme case—point clouds. We propose a tree-based algorithm to reduce the asymptotic complexity of generalized winding number computation, while closely approximating the exact value. Armed with a fast evaluation, we demonstrate the winding number in a variety of new applications: voxelization, signing distances, generating 3D printer paths, defect-tolerant mesh booleans and point set surfaces.

CCS Concepts: • **Computing methodologies** → **Mesh models; Point-based models; Volumetric models; Shape modeling;**

Additional Key Words and Phrases: generalized winding number, inside-outside segmentation, tree-based algorithm, robust geometry processing

## ACM Reference Format:

Gavin Barill, Neil G. Dickson, Ryan Schmidt, David I.W. Levin, and Alec Jacobson. 2018. Fast Winding Numbers for Soups and Clouds. *ACM Trans. Graph.* 37, 4, Article 43 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201337>

# Voxel Cores: Efficient, robust, and provably good approximation of 3D medial axes

YAJIE YAN, Washington University in St. Louis, USA

DAVID LETSCHER, St. Louis University, USA

TAO JU, Washington University in St. Louis, USA

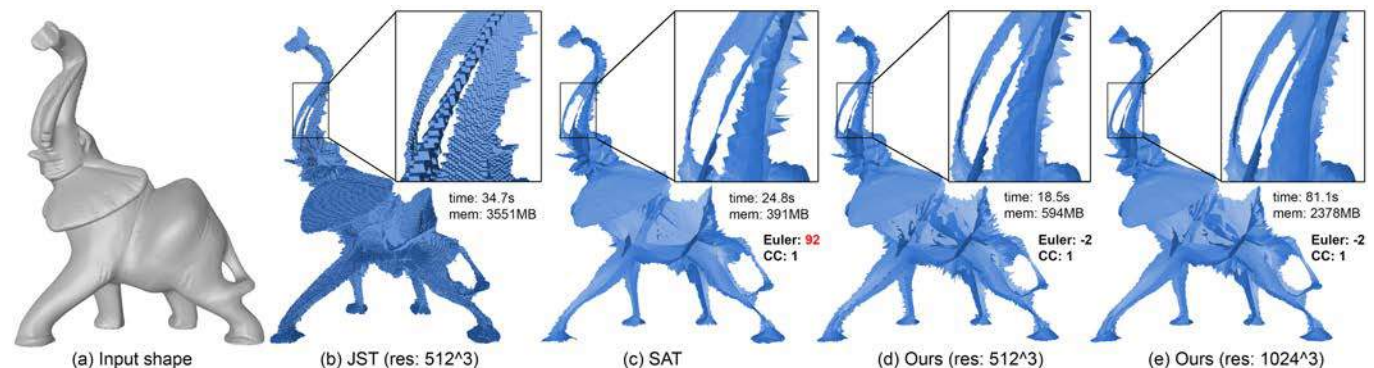


Fig. 1. Approximations of medial axis of the Elephant (a) using the voxel-based method of [Jalba et al. 2016] (noted as JST) (b), the sampling-based method of [Amenta and Kolluri 2001] (as implemented in SAT [Miklos et al. 2010]) (c), and our method at two different voxelization resolutions (d,e). Topological numbers including the Euler characteristic (“Euler”) and number of connected components (“CC”) are noted for methods other than JST. Our method requires significantly less resources than JST, which is unable to handle resolutions higher than  $512^3$  for this input, and produces visually comparable results to SAT but without its topological errors (as evident in the incorrect Euler characteristic).

We present a novel algorithm for computing the medial axes of 3D shapes. We make the observation that the medial axis of a voxel shape can be simply yet faithfully approximated by the interior Voronoi diagram of the boundary vertices, which we call the *voxel core*. We further show that voxel cores can approximate the medial axes of any smooth shape with homotopy equivalence and geometric convergence. These insights motivate an algorithm that is simple, efficient, numerically stable, and equipped with theoretical guarantees. Compared with existing voxel-based methods, our method inherits their simplicity but is more scalable and can process significantly larger inputs. Compared with sampling-based methods that offer similar theoretical guarantees, our method produces visually comparable results but more robustly captures the topology of the input shape.

CCS Concepts: • **Computing methodologies** → **Shape analysis**;

Additional Key Words and Phrases: medial axis, shape analysis, voxelization, Voronoi diagrams

Authors’ addresses: Yajie Yan, Washington University in St. Louis, 1 Brookings Dr. St. Louis, MO, 63130, USA; David Letscher, St. Louis University, 1 N. Grand Blvd. St. Louis, MO, 63130, USA; Tao Ju, Washington University in St. Louis, 1 Brookings Dr. St. Louis, MO, 63130, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART44 \$15.00

<https://doi.org/10.1145/3197517.3201396>

## ACM Reference Format:

Yajie Yan, David Letscher, and Tao Ju. 2018. Voxel Cores: Efficient, robust, and provably good approximation of 3D medial axes. *ACM Trans. Graph.* 37, 4, Article 44 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201396>

## 1 INTRODUCTION

The medial axis [Blum 1967] is a fundamental geometric structure in computer graphics and computer vision. The medial axis  $\mathcal{M}$  of a shape  $O$  is simply defined as the set of points in  $O$  with two or more nearest neighbors on the boundary of  $O$  (see a more formal definition in Section 3.1). The usefulness of the medial axis arises from its many topological and geometric properties. Topologically,  $\mathcal{M}$  is thin (i.e., at least one dimension thinner than  $O$ ) and homotopy equivalent to  $O$ . Geometrically,  $\mathcal{M}$  is centered in  $O$  and captures the protrusions and components of  $O$ . As a result, medial axes have been used in approximating, simplifying, and analyzing shapes. They have also become the foundation for other skeletal shape descriptors [Tagliasacchi et al. 2016].

While simply defined, the medial axis is notoriously difficult to compute. In 3D, the medial axes of most common shapes are complex networks of curved sheets. While methods that compute such networks precisely exist [Culver et al. 2004], they are limited to rather simple shapes due to the high computational cost. To be able to handle complex, real-world data, most existing methods resort to approximations. The goal of approximation is therefore to retain as many properties (both topologically and geometrically) of the medial axis as possible while being able to scale to large inputs.



# Immersion of Self-Intersecting Solids and Surfaces

YIJING LI and JERNEJ BARBIČ, University of Southern California, USA

Self-intersecting, or nearly self-intersecting, meshes are commonly found in 2D and 3D computer graphics practice. Self-intersections occur, for example, in the process of artist manual work, as a by-product of procedural methods for mesh generation, or due to modeling errors introduced by scanning equipment. If the space bounded by such inputs is meshed naively, the resulting mesh joins (“glues”) self-overlapping parts, precluding efficient further modeling and animation of the underlying geometry. Similarly, near self-intersections force the simulation algorithm to employ an unnecessarily detailed mesh to separate the nearly self-intersecting regions. Our work addresses both of these challenges, by giving an algorithm to generate an “un-glued” simulation mesh, of arbitrary user-chosen resolution, that properly accounts for self-intersections and near self-intersections. In order to achieve this result, we study the mathematical concept of immersion, and give a deterministic and constructive algorithm to determine if the input self-intersecting triangle mesh is the boundary of an immersion. For near self-intersections, we give a robust algorithm to properly duplicate mesh elements and correctly embed the underlying geometry into the mesh element copies. Both the self-intersections and near self-intersections are combined into one algorithm that permits successful meshing at arbitrary resolution. Applications of our work include volumetric shape editing, physically based simulation and animation, and volumetric weight and geodesic distance computation on self-intersecting inputs.

CCS Concepts: • **Computing methodologies** → **Volumetric models**; *Physical simulation*;

Additional Key Words and Phrases: self-intersections, meshing, immersion, physically based modeling, animation, computational topology

## ACM Reference Format:

Yijing Li and Jernej Barbič. 2018. Immersion of Self-Intersecting Solids and Surfaces. *ACM Trans. Graph.* 37, 4, Article 45 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201327>

## 1 INTRODUCTION

Polygonal meshes are commonly used in computer graphics due to their versatility and ease of rendering. Many interesting shapes, however, are modeled with geometry that nearly self-intersects, or even self-intersects (Figure 1). Self-intersection may occur simply due to errors in scanning equipment, inadequate modeling by artists; and are sometimes unavoidable. A common example is meshing the mouth of a human face, whereby the upper and lower lip (Figure 2), and/or the mesh of the mouth cavity (teeth and tongue) typically self-intersect in the neutral pose. Another example is meshing the eyes

Authors' address: Yijing Li, [yijingl@usc.edu](mailto:yijingl@usc.edu); Jernej Barbič, [jnb@usc.edu](mailto:jnb@usc.edu), University of Southern California, Computer Science, 941 Bloom Walk SAL 104, Los Angeles, CA, 90089, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART45 \$15.00

<https://doi.org/10.1145/3197517.3201327>

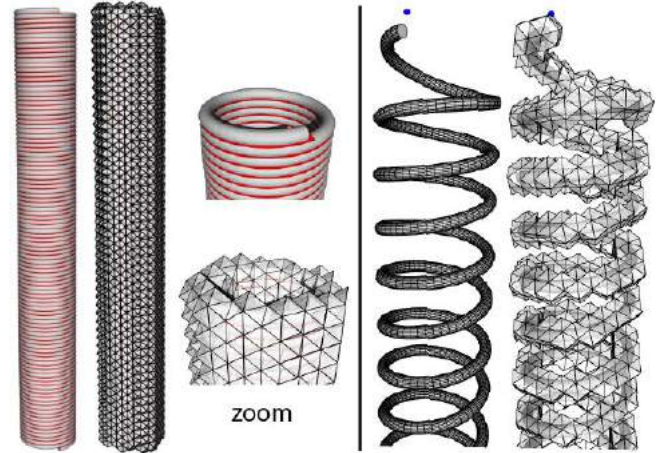


Fig. 1. **Self-intersection-aware tet meshing:** Our method properly duplicated and connected the tets to account for helix triangle mesh self-intersections. Left: input triangle and tetrahedral mesh. Intersecting triangles are shown red. Right: the output tet mesh has been successfully pulled apart using volumetric ARAP [Sorkine and Alexa 2007].

and the eye socket. Self-intersections pose a significant problem for simulation-based modeling and animation pipelines because they preclude an easy generation of a quality simulation mesh. Naive algorithms glue the self-intersecting parts, and/or require extremely small elements to be employed to resolve near self-intersections in the input geometry, resulting in long simulation times.

In our work, we give an algorithm to generate simulation meshes that are properly “un-glued” both for self-intersecting inputs and nearly self-intersecting inputs. Our algorithm is applicable both in 2D and 3D. This is achieved by first generating a world-space simulation mesh (tetrahedral mesh in 3D; triangle mesh in 2D) against the input geometry, using any method and at any user-chosen resolution. Our method then properly duplicates the simulation elements (tets in 3D; triangles in 2D), correctly connects the element copies, and embeds the input geometry into these duplicated elements for correct subsequent modeling and/or animation. Our method works in a unified way both for self-intersecting inputs and nearly self-intersecting inputs. Often, in animation practice, the *only* reason for using detailed meshes is to resolve nearly self-intersecting inputs; *even if the required deformation resolution does not necessitate such detailed meshing*. For nearly self-colliding inputs, inspired by the work of Sifakis [2007], we propose a novel method to unglue tetrahedra containing nearly self-intersecting geometry. We do so using exact arithmetic Sutherland-Hodgman clipping and pseudo-normal tests; which accelerates meshing 30× compared to prior work.

In order to address these intuitive practical goals, one requires a proper mathematical language to express concepts such as “unglued” meshes, “meshing” and “embedding” of overlapping spaces, and state

# Robust Optimization for Topological Surface Reconstruction

ROEE LAZAR, Weizmann Institute of Science

NADAV DYM, Weizmann Institute of Science

YAM KUSHINSKY, Weizmann Institute of Science

ZHIYANG HUANG, Washington University in St. Louis

TAO JU, Washington University in St. Louis

YARON LIPMAN, Weizmann Institute of Science

Surface reconstruction is one of the central problems in computer graphics. Existing research on this problem has primarily focused on improving the geometric aspects of the reconstruction (e.g., smoothness, features, element quality, etc.), and little attention has been paid to ensure it also has desired topological properties (e.g., connectedness and genus). In this paper, we propose a novel and general optimization method for surface reconstruction under topological constraints. The input to our method is a prescribed genus for the reconstructed surface, a partition of the ambient volume into cells, and a set of possible surface candidates and their associated energy within each cell. Our method computes one candidate per cell so that their union is a connected surface with the prescribed genus that minimizes the total energy. We formulate the task as an integer program, and propose a novel solution that combines convex relaxations within a branch and bound framework. As our method is oblivious of the type of input cells, surface candidates, and energy, it can be applied to a variety of reconstruction scenarios, and we explore two of them in the paper: reconstruction from cross-section slices and iso-surfacing an intensity volume. In the first scenario, our method outperforms an existing topology-aware method particularly for complex inputs and higher genus constraints. In the second scenario, we demonstrate the benefit of topology control over classical topology-oblivious methods such as Marching Cubes.

CCS Concepts: • **Computing methodologies** → *Mesh models*;

Additional Key Words and Phrases: surface reconstruction, convex optimization, topological constraints, connectivity constraints

## ACM Reference Format:

Roe Lazar, Nadav Dym, Yam Kushinsky, Zhiyang Huang, Tao Ju, and Yaron Lipman. 2018. Robust Optimization for Topological Surface Reconstruction. *ACM Trans. Graph.* 37, 4, Article 46 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201348>

## 1 INTRODUCTION

Surface reconstruction from incomplete data (e.g., images, volumes, points, curves, etc.) is one of the central topics in computer graphics. To be useful for downstream tasks, the surfaces need to satisfy

Authors' addresses: Roe Lazar, Weizmann Institute of Science; Nadav Dym, Weizmann Institute of Science; Yam Kushinsky, Weizmann Institute of Science; Zhiyang Huang, Washington University in St. Louis; Tao Ju, Washington University in St. Louis; Yaron Lipman, Weizmann Institute of Science.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.

0730-0301/2018/8-ART46 \$15.00

<https://doi.org/10.1145/3197517.3201348>



Fig. 1. Topologically correct, connected, genus zero reconstruction of a corn root (right) from cross sections (left).

application-dependent requirements. One class of such requirements concerns the surfaces' topology, including its connectedness and genus. Topological requirement often arises in biology and medicine, as many natural objects (e.g., anatomical structures) have a known topology; for example the corn roots whose reconstruction is shown in Figure 1 are known to be connected, genus zero surfaces. Geometric processing tasks, such as parameterization and shape matching, also prefer shapes with simple and consistent topology.

Despite the extensive research on surface reconstruction, few works have addressed the topological correctness of their output. The scarcity is in part due to the fact that topology is an inherent property that is invariant of geometric deformations. Hence it is difficult to change topology (without creating adverse impacts on geometry), and even more difficult to enforce a prescribed topology. To satisfy topological requirements, existing reconstruction methods either rely on user interaction, which can be tedious and not suited for batch processing, or resort to post-processing topology repair, which generally has no knowledge of the input data from which the surface is created and hence can make incorrect decisions.

In this work, we propose a solution for automatic topology-aware reconstruction that outputs topologically correct surfaces without the need for post-processing repair. Our method is inspired by the recent works [Huang et al. 2017; Zou et al. 2015] which are among the first topology-aware reconstruction methods. The methods are designed for the problem of surface reconstruction from 2D cross-section curves, and take a divide-and-conquer approach. Given a partitioning of space into "cells" by the cross-section planes, they first enumerate multiple candidate surfaces within each cell that connect the curves on the cell boundary. They then solve a challenging combinatorial optimization problem to select one candidate per cell so that their union surface has the prescribed topology while minimizing a geometric energy. These methods are effective in meeting the topological requirements and are rather efficient for simple data sets consisting of a few planes. However, a simple optimization method (dynamic programming) is used in these methods. Although optimal, dynamic programming scales poorly with the complexity

# Deep Exemplar-based Colorization

MINGMING HE\*, Hong Kong UST

DONGDONG CHEN\*, University of Science and Technology of China

JING LIAO†, Microsoft Research

PEDRO V. SANDER, Hong Kong UST

LU YUAN, Microsoft Research



Fig. 1. Colorization results of black-and-white photographs. Our method provides the capability of generating multiple plausible colorizations by giving different references. Input images (from left to right, top to bottom): Leroy Skalstad/pixabay, Peter van der Sluijs/wikimedia, Bollywood Hungama/wikimedia, Lorri Lang/pixabay, Aamir Mohd Khan/pixabay, Official White House Photographer/wikimedia, Anonymous/wikimedia and K. Krallis/wikimedia.

We propose the first deep learning approach for exemplar-based local colorization. Given a reference color image, our convolutional neural network directly maps a grayscale image to an output colorized image. Rather than using hand-crafted rules as in traditional exemplar-based methods, our end-to-end colorization network learns how to *select*, *propagate*, and *predict* colors from the large-scale data. The approach performs robustly and generalizes well even when using reference images that are unrelated to the input grayscale image. More importantly, as opposed to other learning-based colorization methods, our network allows the user to achieve customizable results by simply feeding different references. In order to further reduce manual effort in selecting the references, the system automatically recommends references with our proposed image retrieval algorithm, which considers both semantic and luminance information. The colorization can be performed fully automatically by simply picking the top reference suggestion.

\*indicates equal contribution.

This work was done when Mingming He and Dongdong Chen were interns at Microsoft Research Asia.

†indicates corresponding author.

Authors' addresses: Mingming He, Department of Computer Science and Engineering, Hong Kong UST, mheah@connect.ust.hk; Dongdong Chen, University of Science and Technology of China, cd722522@mail.ustc.edu.cn; Jing Liao, Microsoft Research, jliao@microsoft.com; Pedro V. Sander, Department of Computer Science and Engineering, Hong Kong UST, psander@cse.ust.hk; Lu Yuan, Microsoft Research, luyuan@microsoft.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART47 \$15.00  
<https://doi.org/10.1145/3197517.3201365>

Our approach is validated through a user study and favorable quantitative comparisons to the state-of-the-art methods. Furthermore, our approach can be naturally extended to video colorization. Our code and models are freely available for public use.

CCS Concepts: • **Computing methodologies** → **Image manipulation**; *Computational photography*; *Neural networks*;

Additional Key Words and Phrases: Colorization, Exemplar-based colorization, Deep learning, Vision for graphics

## ACM Reference Format:

Mingming He\*, Dongdong Chen\*, Jing Liao, Pedro V. Sander and Lu Yuan. 2018. Deep Exemplar-based Colorization. *ACM Trans. Graph.* 37, 4, Article 47 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201365>

## 1 INTRODUCTION

The aim of image colorization is to add colors to a gray image such that the colorized image is perceptually meaningful and visually appealing. The problem is ill-conditioned and inherently ambiguous since there are potentially many colors that can be assigned to the gray pixels of an input image (e.g., leaves may be colored in green, yellow, or brown). Hence, there is no unique correct solution and human intervention often plays an important role in the colorization process.

Manual information to guide the colorization is generally provided in one of two forms: user-guided scribbles or a sample reference image. In the first paradigm [Huang et al. 2005; Levin et al. 2004; Luan et al. 2007; Qu et al. 2006; Yatziv and Sapiro 2006], the manual effort involved in placing the scribbles and the palette of colors must be chosen carefully in order to achieve a convincing result. This often requires both experience and a good sense of aesthetics, thus making it challenging for an untrained user. In the



# Deep Context-Aware Descreening and Rescreening of Halftone Images

TAE-HOON KIM, Intel Corporation

SANG IL PARK, Sejong University, Republic of Korea

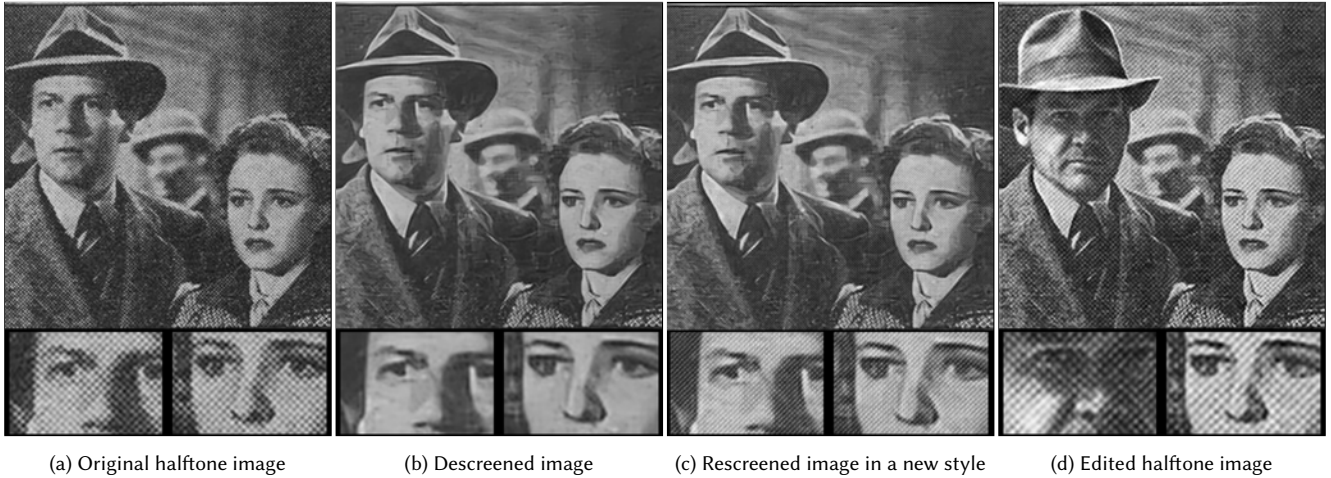


Fig. 1. Descreening and rescreening a halftone image scanned from an old magazine (1940, September 23), *Movie and Radio Guide*. The images below are magnified twice: (a) given original halftone image; (b) descreened image by removing halftone patterns and synthesizing fine details; (c) rescreening the image by applying different halftone styles; (d) editing the descreened image by replacing Joel McCrea's head with Harrison Ford's and applying the original halftone style. Harrison Ford's image adapted from *Indiana Jones and the Last Crusade* (1989) ©Lucasfilm, Ltd.

A fully automatic method for descreening halftone images is presented based on convolutional neural networks with end-to-end learning. Incorporating context level information, the proposed method not only removes halftone artifacts but also synthesizes the fine details lost during halftone. The method consists of two main stages. In the first stage, intrinsic features of the scene are extracted, the low-frequency reconstruction of the image is estimated, and halftone patterns are removed. For the intrinsic features, the edges and object-categories are estimated and fed to the next stage as strong visual and contextual cues. In the second stage, fine details are synthesized on top of the low-frequency output based on an adversarial generative model. In addition, the novel problem of rescreening is addressed, where a natural input image is halftoned so as to be similar to a separately given reference halftone image. To this end, a two-stage convolutional neural network is also presented. Both networks are trained with millions of before-and-after example image pairs of various halftone styles. Qualitative and quantitative evaluations are provided, which demonstrates the effectiveness of the proposed methods.

**CCS Concepts:** • **Computing methodologies** → **Computational photography**;

Authors contributed equally to the work.

Authors' addresses: Tae-hoon Kim, Intel Corporation, [thnkim@gmail.com](mailto:thnkim@gmail.com); Sang Il Park, Sejong University, Republic of Korea, [sipark@sejong.ac.kr](mailto:sipark@sejong.ac.kr).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART48 \$15.00

<https://doi.org/10.1145/3197517.3201377>

**Additional Key Words and Phrases:** computational photography, image processing, halftone, dithering, descreening, image reconstruction, editing

## ACM Reference Format:

Tae-hoon Kim and Sang Il Park. 2018. Deep Context-Aware Descreening and Rescreening of Halftone Images. *ACM Trans. Graph.* 37, 4, Article 48 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201377>

## 1 INTRODUCTION

Since the invention of photography, important personal moments have been recorded by taking photos, and photos of historical importance have been published in newspapers, magazines, and books. These printed photos published several decades ago in mass media can still be easily found in public archives despite the fact that their original films might have been lost. Although invented in the late 19th century, photographic printing techniques have not been fundamentally changed, whereby continuous tone images are simulated using dots of a limited number of colors but of varying size and spacing. These techniques are called *Halftone*. However, owing to the physical limitations of the medium (paper) and the process itself, quality degradation during the halftone process is inevitable; thus, older printed photos exhibit visible dots of non-ignorable size. Therefore, restoring halftone images into high quality images using modern digital technologies is as important as colorizing grayscale photos from the perspective of giving a new breath to old photos.

A typical degradation observed in halftone images is the undesirable regular dot patterns that do not exist in the original. This process also results in sudden tone changes, owing to its limited capacity of tone variations, and in loss of the original details. *Image*

# Non-Stationary Texture Synthesis by Adversarial Expansion

YANG ZHOU\*, Shenzhen University and Huazhong University of Science & Technology

ZHEN ZHU and XIANG BAI, Huazhong University of Science and Technology

DANI LISCHINSKI, The Hebrew University of Jerusalem

DANIEL COHEN-OR, Shenzhen University and Tel Aviv University

HUI HUANG<sup>†</sup>, Shenzhen University



Fig. 1. Examples of two extremely challenging non-stationary textures (middle column), synthesized by our method (left and right). Note that our method succeeds in reproducing and extending the global structure and trends present in the input exemplars.

The real world exhibits an abundance of non-stationary textures. Examples include textures with large scale structures, as well as spatially variant and inhomogeneous textures. While existing example-based texture synthesis methods can cope well with stationary textures, non-stationary textures still pose a considerable challenge, which remains unresolved. In this paper, we propose a new approach for example-based non-stationary texture synthesis. Our approach uses a generative adversarial network (GAN), trained to double the spatial extent of texture blocks extracted from a specific texture exemplar. Once trained, the fully convolutional generator is able to expand the size of the entire exemplar, as well as of any of its sub-blocks. We demonstrate that this conceptually simple approach is highly effective for capturing large scale structures, as well as other non-stationary attributes of the input exemplar. As a result, it can cope with challenging textures, which, to our knowledge, no other existing method can handle.

CCS Concepts: • **Computing methodologies** → **Appearance and texture representations**; *Image manipulation*; *Texturing*;

\*Yang Zhou and Zhen Zhu are joint first authors

<sup>†</sup>Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Yang Zhou, Shenzhen University, Huazhong University of Science & Technology; Zhen Zhu, Xiang Bai, Huazhong University of Science and Technology; Dani Lischinski, The Hebrew University of Jerusalem; Daniel Cohen-Or, Shenzhen University, Tel Aviv University; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART49 \$15.00

<https://doi.org/10.1145/3197517.3201286>

Additional Key Words and Phrases: Example-based texture synthesis, non-stationary textures, generative adversarial networks

## ACM Reference Format:

Yang Zhou, Zhen Zhu, Xiang Bai, Dani Lischinski, Daniel Cohen-Or, and Hui Huang. 2018. Non-Stationary Texture Synthesis by Adversarial Expansion. *ACM Trans. Graph.* 37, 4, Article 49 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201286>

## 1 INTRODUCTION

Example-based texture synthesis has been an active area of research for over two decades. However, despite excellent results for many classes of textures, example-based synthesis of significantly non-stationary textures remains an open problem. Examples of non-stationary textures include textures with large-scale irregular structures, or ones that exhibit spatial variance in various attributes, such as color, local orientation, and local scale. Inhomogeneous textures, such as weathered surfaces are another challenging example of non-stationarity. Several challenging non-stationary examples are shown in Figures 1 and 2.

During the previous decade, stitching-based [Efros and Freeman 2001; Kwatra et al. 2003] and optimization-based [Kwatra et al. 2005; Wexler et al. 2007] approaches have established themselves as highly effective for example-based texture synthesis. More recently, deep learning based approaches for texture synthesis have begun to gain popularity. However, Figure 7 demonstrates that none of the existing state-of-the-art methods are able to successfully cope with significantly non-stationary input exemplars. Depending on the assumptions of each method, the results are either stationary or periodic, failing to mimic the large-scale structure and spatially variant attributes of the exemplars.

# Eulerian-on-Lagrangian Cloth Simulation

NICHOLAS J. WEIDNER, Texas A&M University

KYLE PIDDINGTON, California Polytechnic State University

DAVID I.W. LEVIN, The University of Toronto

SHINJIRO SUEDA, Texas A&M University

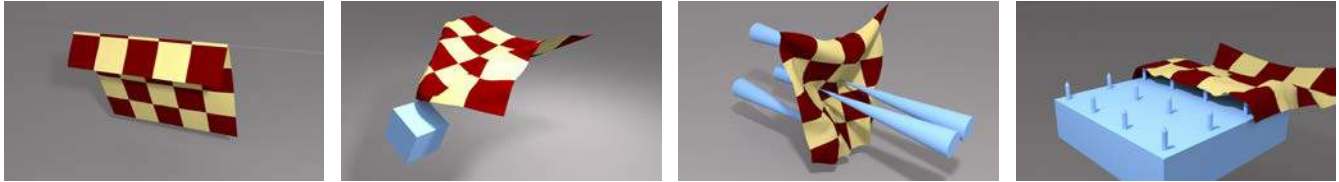


Fig. 1. Still frames from our Eulerian-on-Lagrangian cloth simulations. With our framework, we are able to simulate the smooth sliding of cloth over sharp features such as wires, moving boxes, sharp teeth, and nails.

We resolve the longstanding problem of simulating the contact-mediated interaction of cloth and sharp geometric features by introducing an Eulerian-on-Lagrangian (EOL) approach to cloth simulation. Unlike traditional Lagrangian approaches to cloth simulation, our EOL approach permits bending exactly at and sliding over sharp edges, avoiding parasitic locking caused by over-constraining contact constraints. Wherever the cloth is in contact with sharp features, we insert EOL vertices into the cloth, while the rest of the cloth is simulated in the standard Lagrangian fashion. Our algorithm manifests as new equations of motion for EOL vertices, a contact-conforming remesh, and a set of simple constraint assignment rules, all of which can be incorporated into existing state-of-the-art cloth simulators to enable smooth, inequality-constrained contact between cloth and objects in the world.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: Physical simulation, Lagrangian mechanics, Cloth, Eulerian, Constraints, Contact

## ACM Reference Format:

Nicholas J. Weidner, Kyle Piddington, David I.W. Levin, and Shinjiro Sueda. 2018. Eulerian-on-Lagrangian Cloth Simulation. *ACM Trans. Graph.* 37, 4, Article 50 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201281>

## 1 INTRODUCTION

Cloth simulation has a long history in computer graphics. Starting with the work by Terzopoulos et al. [1987], cloth simulation algorithms have steadily evolved from research curiosities to an integral

part of the visual effects pipeline. In fact, cloth simulation has proven to be one of the major success stories for physics-based animation as its phenomenological behavior is difficult to capture for an artist and far simpler to describe using the laws of physics. This allows artist-controlled inputs, such as geometry and initial conditions, to generate life-like results with much less human effort. Motivated by this, research over the last few decades has focused on improving both the performance and physical realism of cloth-simulation.

Due to its highly-deformable nature, cloth undergoes complex colliding interactions with other geometry and itself. Therefore, it is no surprise that much of the research in the field focuses on resolving these collisions. Despite this concentration of effort, there is one crucial scenario which, so far, has yet to be treated: the interaction of cloth with sharp geometric features. Such interactions are common when cloth interacts with the world, from a table cloth pulled over a table edge, to a sheet dragged off a clothing line to the unveiling of a sculpture. Yet this literal (but not figurative) edge case still baffles today's state-of-the-art approaches.

The difficulty arises because all previous cloth simulation algorithms rely on a Lagrangian discretization of the cloth, which only permits bending along its edges. As a pedagogical example, imagine draping the cloth over a sharp edge of a table. Unless the cloth mesh has edges exactly aligned with the table edge, the cloth will not be able to bend sharply, leading to unappealing visual artifacts. For a stationary piece of cloth, this can be resolved by remeshing the cloth, inserting edges that directly align with the table edge. However, what happens if we pull the cloth? We suffer from unacceptable jitter since the cloth can only be remeshed between time steps. This occurs even if we remesh the cloth during the sliding motion, as shown in Fig. 2. In certain scenarios we may even experience catastrophic locking, wherein the cloth, unable to slide over the edge, simply gets stuck. Today, standard approaches to fixing this problem involve perturbing the underlying geometry by approximating the underlying surface as smooth (e.g., by averaging surface normals). However, this approach is unsatisfactory because it prevents us from visually capturing the correct physical sliding behavior of the cloth over the sharp feature, even if this is precisely what an artist wants to achieve. This goes against the very philosophy of physics-based

Authors' addresses: Nicholas J. Weidner, Department of Computer Science & Engineering, Texas A&M University, College Station, TX, [weidnerjn@tamu.edu](mailto:weidnerjn@tamu.edu); Kyle Piddington, Department of Computer Science & Software Engineering, California Polytechnic State University, San Luis Obispo, CA, [kpiddy@gmail.com](mailto:kpiddy@gmail.com); David I.W. Levin, Department of Computer Science, The University of Toronto, Toronto, ON, [diwlevin@cs.toronto.edu](mailto:diwlevin@cs.toronto.edu); Shinjiro Sueda, Department of Computer Science & Engineering, Texas A&M University, College Station, TX, [sueda@tamu.edu](mailto:sueda@tamu.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART50 \$15.00

<https://doi.org/10.1145/3197517.3201281>



# A Multi-Scale Model for Simulating Liquid-Fabric Interactions

YUN (RAYMOND) FEI, Columbia University, USA

CHRISTOPHER BATTY, University of Waterloo, Canada

EITAN GRINSPUN and CHANGXI ZHENG, Columbia University, USA

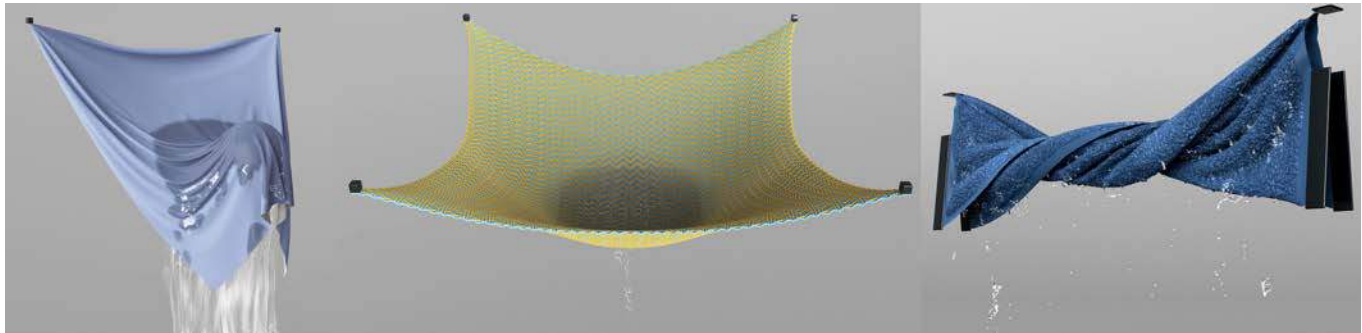


Fig. 1. **Left:** A piece of mesh-based cloth draped over a solid obstacle is splashed with water. **Middle:** Water flows through a piece of yarn-based handwoven fabric. **Right:** After being vigorously wrung out, a fuzzy towel continues to drip.

We propose a method for simulating the complex dynamics of partially and fully saturated woven and knit fabrics interacting with liquid, including the effects of buoyancy, nonlinear drag, pore (capillary) pressure, dripping, and convection-diffusion. Our model evolves the velocity fields of both the liquid and solid relying on mixture theory, as well as tracking a scalar saturation variable that affects the pore pressure forces in the fluid. We consider the porous microstructure implied by the fibers composing individual threads, and use it to derive homogenized drag and pore pressure models that faithfully reflect the anisotropy of fabrics. In addition to the bulk liquid and fabric motion, we derive a quasi-static flow model that accounts for liquid spreading within the fabric itself. Our implementation significantly extends standard numerical cloth and fluid models to support the diverse behaviors of wet fabric, and includes a numerical method tailored to cope with the challenging nonlinearities of the problem. We explore a range of fabric-water interactions to validate our model, including challenging animation scenarios involving splashing, wringing, and collisions with obstacles, along with qualitative comparisons against simple physical experiments.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: wet cloth, wet yarn, fluid dynamics, two-way coupling, multi-scale

Authors' addresses: Yun (Raymond) Fei, Columbia University, Computer Science, New York, NY, 10027, USA; Christopher Batty, University of Waterloo, Computer Science, Waterloo, ON, N2L 3G1, Canada; Eitan Grinspun; Changxi Zheng, Columbia University, Computer Science, New York, NY, 10027, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART51 \$15.00

<https://doi.org/10.1145/3197517.3201392>

## ACM Reference Format:

Yun (Raymond) Fei, Christopher Batty, Eitan Grinspun, and Changxi Zheng. 2018. A Multi-Scale Model for Simulating Liquid-Fabric Interactions. *ACM Trans. Graph.* 37, 4, Article 51 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201392>

## 1 INTRODUCTION

A beach vacation provides ample opportunity to experience the characteristic aspects of liquid-fabric interaction. A tipped piña colada splashes onto beachwear, wetting, then diffusing to dampen a larger area. Submerged board shorts drag along with the ocean waves, lifted buoyantly upon the surf, then drip distinctively on the return to dry land. Wringing those shorts then squeezes out the liquid, leaving tiny drops scattered within the fabric microstructure.

To develop a computational model of these varied liquid-fabric interactions, we must understand the composition of fabric. Fabric is composed of individual strands (“thread,” “yarn”) packed into thin oriented fibers. Tiny pockets within and between these fibers collect fluid, and are largely responsible for the wetting behavior we observe at the coarse scale. Because these pockets are numerous and individually imperceptible to the naked eye, it can be wasteful or intractable to represent them as discrete elements for animation applications. Therefore, we develop a macroscopic model.

Building on modern mixture theory [Anderson and Jackson 1967], we model fabric as a continuous porous medium through which fluid may flow. The model accounts for the material’s anisotropic structure, and the evolution of its saturation, to capture buoyancy, drag, small-scale capillary (surface tension) effects, and fluid convection.

Our numerical treatment integrates a piecewise linear Lagrangian cloth or rod model [Bergou et al. 2010; Grinspun et al. 2003] with a hybrid Eulerian-Lagrangian (APIC) fluid simulator [Bridson 2015; Jiang et al. 2015]. We apply this model to application scenarios involving mesh-based cloth, yarn-based fabric, and fuzzy fabric

# An Implicit Frictional Contact Solver for Adaptive Cloth Simulation

JIE LI, University of Minnesota, USA

GILLES DAVIET, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France and Weta Digital, New Zealand

RAHUL NARAIN, University of Minnesota, USA and Indian Institute of Technology Delhi, India

FLORENCE BERTAILS-DESCOUBES, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France

MATTHEW OVERBY and GEORGE E. BROWN, University of Minnesota, USA

LAURENCE BOISSIEUX, Univ. Grenoble Alpes, Inria, France



Fig. 1. Our implicit solver simultaneously resolves cloth elasticity, non-penetration and exact Coulomb friction constraints at every body-cloth and cloth-cloth contact, largely improving physical realism over previous methods. This new solver especially allows us to simulate accurately the effect of a varying friction coefficient  $\mu$ , capturing a diversity of cloth sliding motions and folding patterns as shown in this batwing dress example (from left to right,  $\mu = 0$ ,  $\mu = 0.1$ ,  $\mu = 0.3$ , and  $\mu = 0.6$ ). In this example featuring 2,600 contact points on average, our solver converges at each time step ( $dt = 2ms$ ) to a high precision in a few hundred milliseconds only.

Cloth dynamics plays an important role in the visual appearance of moving characters. Properly accounting for contact and friction is of utmost importance to avoid cloth-body and cloth-cloth penetration and to capture typical folding and stick-slip behavior due to dry friction. We present here the first method able to account for cloth contact with exact Coulomb friction, treating both cloth self-contacts and contacts occurring between the cloth and an underlying character. Our key contribution is to observe that for a nodal system like cloth, the frictional contact problem may be formulated based on velocities as primary variables, without having to compute the costly Delassus operator. Then, by reversing the roles classically played by the velocities

and the contact impulses, conical complementarity solvers of the literature can be adapted to solve for compatible velocities at nodes. To handle the full complexity of cloth dynamics scenarios, we have extended this base algorithm in two ways: first, towards the accurate treatment of frictional contact at any location of the cloth, through an adaptive node refinement strategy; second, towards the handling of multiple constraints at each node, through the duplication of constrained nodes and the adding of pin constraints between duplicata. Our method allows us to handle the complex cloth-cloth and cloth-body interactions in full-size garments with an unprecedented level of realism compared to former methods, while maintaining reasonable computational timings.

Authors' addresses: Jie Li, lixx4611@umn.edu, University of Minnesota, USA; Gilles Daviet, gdaviet@wetafx.co.nz, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000, Grenoble, France, Weta Digital, New Zealand; Rahul Narain, narain@cse.iitd.ac.in, University of Minnesota, USA, Indian Institute of Technology Delhi, India; Florence Bertails-Descoubes, florence.descoubes@inria.fr, Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000, Grenoble, France; Matthew Overby, over0219@umn.edu; George E. Brown, brow2327@umn.edu, University of Minnesota, USA; Laurence Boissieux, laurence.boissieux@inria.fr, Univ. Grenoble Alpes, Inria, 38000, Grenoble, France.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART52 \$15.00

<https://doi.org/10.1145/3197517.3201308>

CCS Concepts: • **Computing methodologies** → **Animation; Physical simulation**;

Additional Key Words and Phrases: cloth simulation, dry frictional contact, implicit solver

## ACM Reference Format:

Jie Li, Gilles Daviet, Rahul Narain, Florence Bertails-Descoubes, Matthew Overby, George E. Brown, and Laurence Boissieux. 2018. An Implicit Frictional Contact Solver for Adaptive Cloth Simulation. *ACM Trans. Graph.* 37, 4, Article 52 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201308>

## 1 INTRODUCTION

Cloth simulation has been a very active research area over the last decades, yielding compelling animations of cloth in feature film movies that are less and less distinguishable from reality. Relying upon mature cloth simulators, a myriad of new applications has

# Rule-Free Sewing Pattern Adjustment with Precision and Efficiency

HUAMIN WANG, The Ohio State University and Frilly Inc.

Being able to customize sewing patterns for different human bodies without using any pre-defined adjustment rule will not only improve the realism of virtual humans in the entertainment industry, but also deeply affect the fashion industry by making fast fashion and made-to-measure garments more accessible. To meet the requirement set by the fashion industry, a sewing pattern adjustment system must be both efficient and precise, which unfortunately cannot be achieved by existing techniques. In this paper, we propose to solve sewing pattern adjustment as a nonlinear optimization problem immediately, rather than in two phases: a garment shape optimization phase and an inverse pattern design phase as in previous systems. This allows us to directly minimize the objective function that evaluates the fitting quality of the garment sewn from a pattern, without any compromise caused by the nonexistence of the solution to inverse pattern design. To improve the efficiency of our system, we carry out systematic research on a variety of optimization topics, including pattern parametrization, initialization, an inexact strategy, acceleration, and CPU-GPU implementation. We verify the usability of our system through automatic grading tests and made-to-measure tests. Designers and pattern makers confirm that our pattern results are able to preserve design details and their fitting qualities are acceptable. In our computational experiment, the system further demonstrates its efficiency, reliability, and flexibility of handling various pattern designs. While our current system still needs to overcome certain limitations, we believe it is a crucial step toward fully automatic pattern design and adjustment in the future.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: sewing pattern, made-to-measure, nonlinear optimization, quasistatic simulation, inexact methods

## ACM Reference Format:

Huamin Wang. 2018. Rule-Free Sewing Pattern Adjustment with Precision and Efficiency. *ACM Trans. Graph.* 37, 4, Article 53 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201320>

## 1 INTRODUCTION

Today the multi-billion fashion industry is dominated by *ready-to-wear* garments, made and sold in standardized sizes, from Size 00 to Size 12 for example. To make a ready-to-wear garment, a pattern maker typically develops a sewing pattern in a base size first, such as Size 6. A pattern grader then transforms the pattern from the base size to additional sizes, using a series of *garment grading rules*



(a) The sewing patterns in three standardized sizes



(b) Simulation (size 2)

(c) Simulation (size 6)

(d) Simulation (size 12)



(e) Real dress (size 2)

(f) Real dress (size 6)

(g) Real dress (size 12)

Fig. 1. The grading results of a secretary dress example. Given a base sewing pattern (black) in (a) and its simulated garment in (c), our pattern optimization system automatically converts the base pattern into new patterns for making garments that fit new bodies, as shown in (b) and (d). The experiment demonstrates that the real garments in (e) and (g) are able to preserve design details, and the whole optimization process lasts less than two minutes for a mesh with 66K elements.

Author's address: Huamin Wang, Department of Computer Science and Engineering, The Ohio State University and Frilly Inc. [whmin@cse.ohio-state.edu](mailto:whmin@cse.ohio-state.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART53 \$15.00

<https://doi.org/10.1145/3197517.3201320>

established over the last two centuries. Compared with a ready-to-wear garment, a *made-to-measure* garment reduces the inventory cost and provides superior fit, by customizing an existing pattern for each specific individual as needed. However, a made-to-measure garment is much more difficult and time consuming to develop, since the individual's body can differ significantly and the *pattern adjustment rules* for disproportionate body parts are rather sparse,



# Analytic Spherical Harmonic Coefficients for Polygonal Area Lights

JINGWEN WANG, University of California, San Diego

RAVI RAMAMOORTHY, University of California, San Diego

Spherical Harmonic (SH) lighting is widely used for real-time rendering within Precomputed Radiance Transfer (PRT) systems. SH coefficients are precomputed and stored at object vertices, and combined interactively with SH lighting coefficients to enable effects like soft shadows, interreflections, and glossy reflection. However, the most common PRT techniques assume distant, low-frequency environment lighting, for which SH lighting coefficients can easily be computed once per frame. There is currently limited support for near-field illumination and area lights, since it is non-trivial to compute the SH coefficients for an area light, and the incident lighting (SH coefficients) varies over the object geometry. We present an efficient closed-form solution for projection of uniform polygonal area lights to spherical harmonic coefficients of arbitrary order, enabling easy adoption of accurate area lighting in PRT systems, with no modifications required to the core PRT framework. Our method only requires computing zonal harmonic (ZH) coefficients, for which we introduce a novel recurrence relation. In practice, ZH coefficients are built up iteratively, with computation linear in the desired SH order. General SH coefficients can then be obtained by the recently developed sparse zonal harmonic rotation method.

CCS Concepts: • **Computing Methodologies** → **Computer Graphics**; **Rendering**;

Additional Key Words and Phrases: PRT, spherical harmonics, area lighting

## ACM Reference Format:

Jingwen Wang and Ravi Ramamoorthi. 2018. Analytic Spherical Harmonic Coefficients for Polygonal Area Lights. *ACM Trans. Graph.* 37, 4, Article 54 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201291>

## 1 INTRODUCTION

Spherical Harmonic (SH) lighting is a popular method in interactive rendering, within Precomputed Radiance Transfer (PRT) systems [Sloan et al. 2002]. While strictly accurate only for low-frequency lighting, the method enables realistic visual effects like soft shadows and glossy reflections at real-time frame rates with dynamic lighting. PRT has been the subject of a substantial body of work (see survey [Ramamoorthi 2009]), and is widely used in real-time applications like video games.<sup>1</sup> In recent years, SH lighting and PRT have also become mainstays of offline rendering, and have been widely used in movie production [Pantaleoni et al. 2010].

<sup>1</sup>PRT was at one time part of the DirectX API; modern games use a variety of real-time techniques, including SH irradiance, volumetric transport, and variants of PRT.

Authors' addresses: Jingwen Wang, University of California, San Diego, [jiw524@eng.ucsd.edu](mailto:jiw524@eng.ucsd.edu); Ravi Ramamoorthi, University of California, San Diego, [ravir@cs.ucsd.edu](mailto:ravir@cs.ucsd.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART54 \$15.00  
<https://doi.org/10.1145/3197517.3201291>

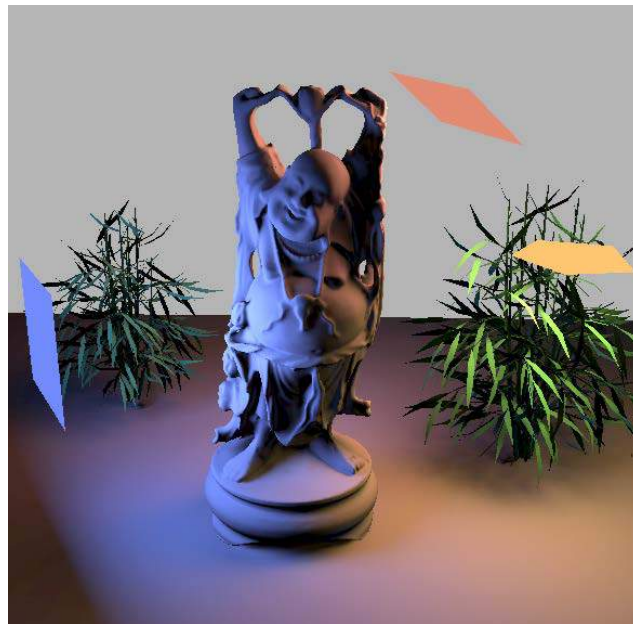


Fig. 1. Images rendered in a PRT system with multiple polygonal area lights, with analytic SH lighting coefficients computed in real-time with our method. This scene contains 388K triangles and is rendered at 28.4 frames per second on an NVIDIA GeForce GTX 1080 Ti GPU.

The core PRT framework precomputes a transfer function for each vertex on object geometry, taking into account visibility and BRDF effects. This transfer function is then converted (projected) into spherical harmonics and stored on the object, possibly with compression [Sloan et al. 2003]. For soft shadows on a Lambertian surface, this would simply be an SH representation of the visibility in each direction. At run-time, these SH coefficients are combined with the lighting (a simple dot product of SH coefficients of lighting and transfer for visibility). In this paper, we do not modify the pre-computation or core PRT framework at all, but address computation of the SH lighting coefficients. Most PRT methods assume distant, low-frequency environment map lighting. In this case, the incident lighting is the same everywhere in the scene, and SH lighting coefficients can easily be computed once per frame using standard numerical integration, and re-used for each vertex on the object.

On the other hand, there is currently limited support for near-field illumination from area light sources. Indeed, near-field effects imply that the angular distribution of incident radiance is now different at each vertex (location in space), and numerical integration to compute SH lighting coefficients separately for each vertex or pixel is too expensive. In this paper, we address this challenge by developing an efficient closed-form computation for spherical harmonic coefficients of general orders, for an arbitrary polygonal area light. As in most previous work, we assume the light source emission is

# Laplacian Kernel Splatting for Efficient Depth-of-field and Motion Blur Synthesis or Reconstruction

THOMAS LEIMKÜHLER, MPI Informatik, Saarland Informatics Campus, Germany

HANS-PETER SEIDEL, MPI Informatik, Saarland Informatics Campus, Germany

TOBIAS RITSCHER, University College London, United Kingdom

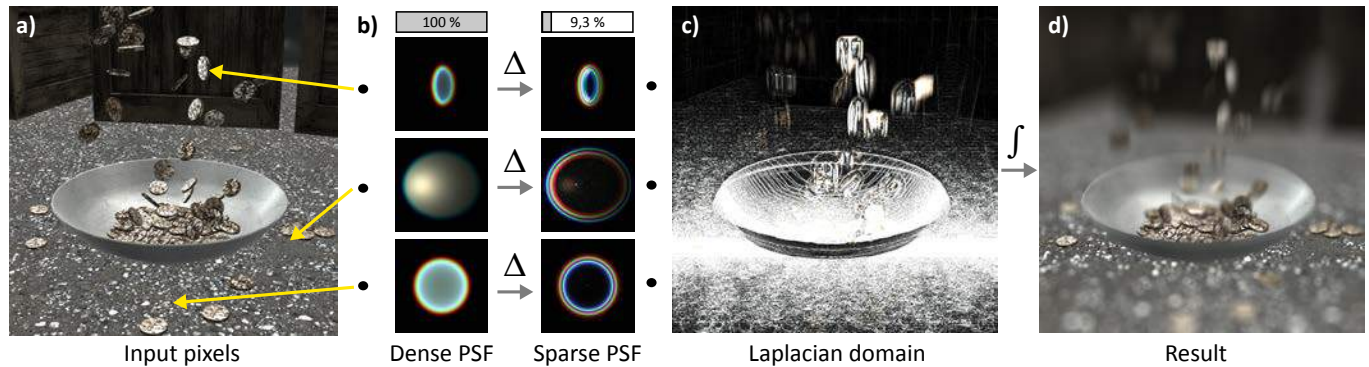


Fig. 1. Computing motion blur and depth-of-field by applying a point spread function (PSF) to every pixel (a) is computationally costly. We suggest splatting a pre-computed sparse approximation of the Laplacian of a PSF (b) to the Laplacian of an image (c) that under integration provides the same result (d). Note the circular bokeh combined with motion blur (1024×1024 pixels, 2 layers, 190 ms, Nvidia GTX 980Ti at .97 SSIM to a path-traced reference).

Simulating combinations of depth-of-field and motion blur is an important factor to cinematic quality in synthetic images but can take long to compute. Splatting the point-spread function (PSF) of every pixel is general and provides high quality, but requires prohibitive compute time. We accelerate this in two steps: In a pre-process we optimize for sparse representations of the Laplacian of all possible PSFs that we call *spreadlets*. At runtime, spreadlets can be splat efficiently to the Laplacian of an image. Integrating this image produces the final result. Our approach scales faithfully to strong motion and large out-of-focus areas and compares favorably in speed and quality with off-line and interactive approaches. It is applicable to both synthesizing from pinhole as well as reconstructing from stochastic images, with or without layering.

CCS Concepts: • **Computing methodologies** → **Rasterization; Image-based rendering**; *Massively parallel algorithms*;

Additional Key Words and Phrases: Distribution rendering; Gradient Domain; Depth-of-field; Motion blur.

Authors' addresses: Thomas Leimkühler, MPI Informatik, Saarland Informatics Campus, Campus E1.4, Saarbrücken, 66123, Germany, tleimkueh@mpi-inf.mpg.de; Hans-Peter Seidel, MPI Informatik, Saarland Informatics Campus, Saarbrücken, Germany; Tobias Ritschel, University College London, London, United Kingdom.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://doi.org/10.1145/3197517.3201379).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART55 \$15.00

<https://doi.org/10.1145/3197517.3201379>

## ACM Reference Format:

Thomas Leimkühler, Hans-Peter Seidel, and Tobias Ritschel. 2018. Laplacian Kernel Splatting for Efficient Depth-of-field and Motion Blur Synthesis or Reconstruction. *ACM Trans. Graph.* 37, 4, Article 55 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201379>

## 1 INTRODUCTION

Depth-of-field (DoF) and motion blur (MB) are a key ingredient to the look and feel of most cinematic-quality feature films [Goy 2013]. Reproducing them in synthesized imagery is a typical and well-understood part of most photo-realistic rendering systems. When it comes to efficient, interactive or even real-time rendering, current solutions to DoF and MB typically make several key assumptions that result in computational efficiency but come at the cost of reduced quality. A typical example is to assume DoF and MB to be independent, to be able to approximate the space-time lens transport by a convolution [Potmesil and Chakravarty 1981] and often to approximate their reconstruction using Gaussian filtering [Belcour et al. 2013; Egan et al. 2009; Munkberg et al. 2014; Soler et al. 2009; Vaidyanathan et al. 2015]. In this work, we devise a method to synthesize or reconstruct cinematic quality motion blur and depth-of-field, while retaining most of the efficiency of typical approximations.

Input to our method are pixels (Fig. 1, a), which we see as light-field samples, labeled with additional geometric and dynamic information. We can work on pixels coming both from simple and layered images, as well as from a pinhole camera (synthesis), or stochastic path-tracing/rasterization (reconstruction). The point spread function (PSF) of each input pixel can affect a very large image area.

# Deep Learning of Biomimetic Sensorimotor Control for Biomechanical Human Animation

MASAKI NAKADA, TAO ZHOU, HONGLIN CHEN, TOMER WEISS, and DEMETRI TERZOPOULOS,  
University of California, Los Angeles, USA

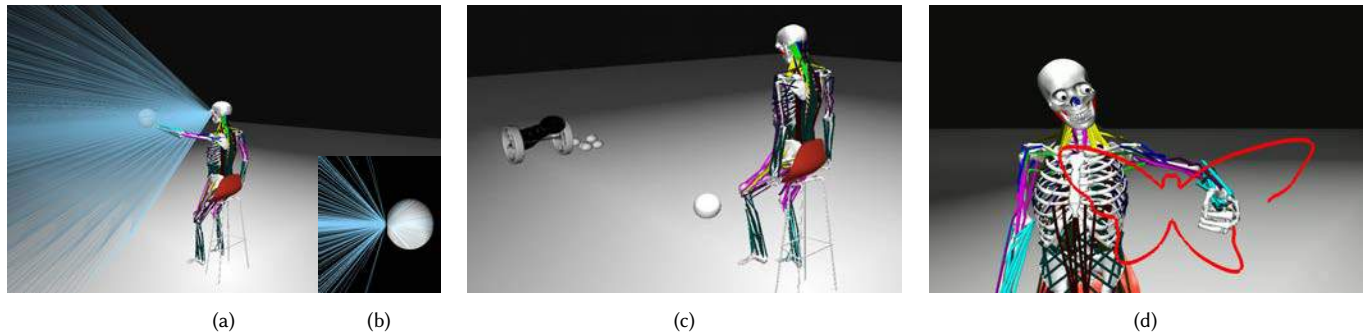


Fig. 1. Our anatomically accurate biomechanical human model features a biomimetic sensorimotor system that incorporates 20 trained deep neural networks, 10 devoted to visual perception and 10 to motor control. Its motor subsystem comprises 5 neuromuscular controllers that actuate the cervicocephalic and four limb musculoskeletal complexes. Via ray tracing (blue rays), its virtual eyes (b) visually sample the environment (a) to compute the irradiance at RGB photoreceptors nonuniformly distributed on their foveated retinas. Its sensory subsystem performs visual analysis that drives the motor subsystem to actuate appropriate eye, head, and limb movements. Our virtual human can visually foveate and track a moving ball, and intercept it with arms (a) and legs (c), as well as write and draw with its finger (d). The rendered appearance of the numerous Hill-type skeletal muscles is unimportant in our work; they are depicted as lines, color coded to distinguish the different muscle groups, with brightness proportional to each contractile muscle's efferent neural activation.

We introduce a biomimetic framework for human sensorimotor control, which features a biomechanically simulated human musculoskeletal model actuated by numerous muscles, with eyes whose retinas have nonuniformly distributed photoreceptors. The virtual human's sensorimotor control system comprises 20 trained deep neural networks (DNNs), half constituting the neuromuscular motor subsystem, while the other half compose the visual sensory subsystem. Directly from the photoreceptor responses, 2 vision DNNs drive eye and head movements, while 8 vision DNNs extract visual information required to direct arm and leg actions. Ten DNNs achieve neuromuscular control—2 DNNs control the 216 neck muscles that actuate the cervicocephalic musculoskeletal complex to produce natural head movements, and 2 DNNs control each limb; i.e., the 29 muscles of each arm and 39 muscles of each leg. By synthesizing its own training data, our virtual human automatically learns efficient, online, active visuomotor control of its eyes, head, and limbs in order to perform nontrivial tasks involving the foveation and visual pursuit of target objects coupled with visually-guided limb-reaching actions to intercept the moving targets, as well as to carry out drawing and writing tasks.

Authors' address: Masaki Nakada, nakada@cs.ucla.edu; Tao Zhou, taozhou@cs.ucla.edu; Honglin Chen, chenhonglin@g.ucla.edu; Tomer Weiss, tweiss@cs.ucla.edu; Demetri Terzopoulos, dt@cs.ucla.edu, Computer Science Department, University of California, Los Angeles, Los Angeles, CA, 90095, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART56 \$15.00  
<https://doi.org/10.1145/3197517.3201305>

CCS Concepts: • **Computing methodologies** → **Computer graphics; Animation; Bio-inspired approaches; Neural networks; Artificial life; Physical simulation**;

Additional Key Words and Phrases: Biomechanical human animation; Sensorimotor control; Biomimetic perception; Deep neural network learning.

## ACM Reference Format:

Masaki Nakada, Tao Zhou, Honglin Chen, Tomer Weiss, and Demetri Terzopoulos. 2018. Deep Learning of Biomimetic Sensorimotor Control for Biomechanical Human Animation. *ACM Trans. Graph.* 37, 4, Article 56 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201305>

## 1 INTRODUCTION

The modeling of graphical characters based on human anatomy is attracting much attention in computer animation. Increasingly accurate biomechanical modeling of the human body should, in principle, yield more realistic human animation. However, it remains a daunting challenge to control complex biomechanical human musculoskeletal models with over two hundred bones and on the order of a thousand contractile muscle actuators. This paper reports on a major stride toward achieving this goal. We take a neuromuscular control approach based on machine learning using Deep Neural Networks (DNNs). Beyond pure motor control, we confront the difficult, less explored problem of sensorimotor control, in which processed sensory information acquired through active perception drives appropriate motor responses that support natural behavior.

More specifically, the main research problems that we tackle in this paper are as follows:



# Dexterous Manipulation and Control with Volumetric Muscles

SEUNGHWAN LEE, Seoul National University  
RI YU, Seoul National University  
JUNGNAM PARK, Seoul National University  
MRIDUL AANJANEYA, Rutgers University  
EFTYCHIOS SIFAKIS, University of Wisconsin-Madison  
JEHEE LEE, Seoul National University

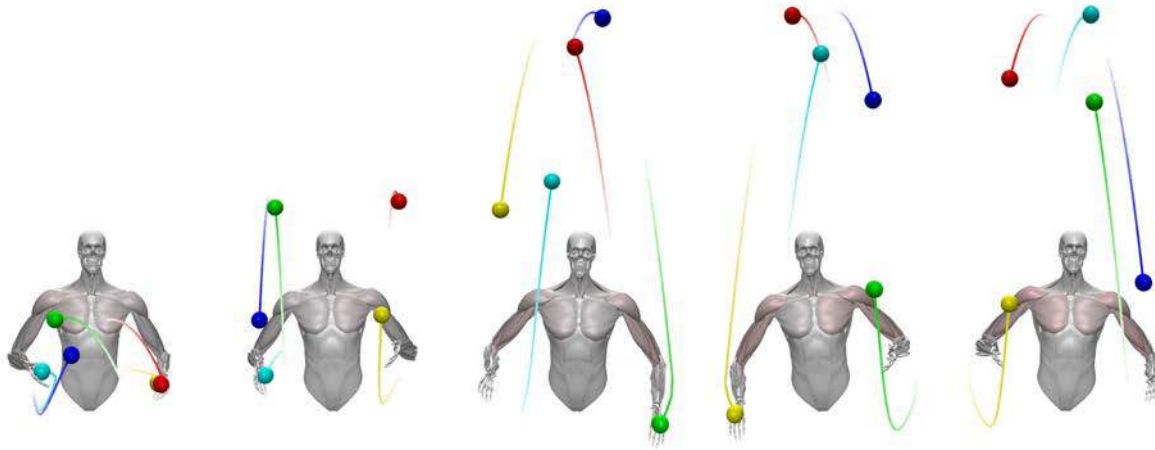


Fig. 1. Demonstration of motion control in an anatomical simulation of a juggling task, actuated by volumetric muscles. Active muscles are shown in pink.

We propose a framework for simulation and control of the human musculoskeletal system, capable of reproducing realistic animations of dexterous activities with high-level coordination. We present the first controllable system in this class that incorporates volumetric muscle actuators, tightly coupled with the motion controller, in enhancement of line-segment approximations that prior art is overwhelmingly restricted to. The theoretical framework put forth by our methodology computes all the necessary Jacobians for control, even with the drastically increased dimensionality of the state descriptors associated with three-dimensional, volumetric muscles. The direct coupling of volumetric actuators in the controller allows us to model muscular deficiencies that manifest in shape and geometry, in ways

that cannot be captured with line-segment approximations. Our controller is coupled with a trajectory optimization framework, and its efficacy is demonstrated in complex motion tasks such as juggling, and weightlifting sequences with variable anatomic parameters and interaction constraints.

CCS Concepts: • **Computing methodologies** → **Physical simulation**; **Motion processing**;

Additional Key Words and Phrases: Motion control, Volumetric Muscles, Musculoskeletal simulation, Trajectory Optimization

## ACM Reference Format:

Seunghwan Lee, Ri Yu, Jungnam Park, Mridul Aanjaneya, Eftychios Sifakis, and Jehee Lee. 2018. Dexterous Manipulation and Control with Volumetric Muscles. *ACM Trans. Graph.* 37, 4, Article 57 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201330>

## 1 INTRODUCTION

Complex movement of a human body emerges from the geometrical structure of the musculoskeletal system and its mechanical characteristics. The skeleton supports the body, and the muscles contract to induce motion of the attached bones. The human brain coordinates the activations of the muscles in harmonious synergy, to create intended bulk motion or maintain kinematic balance. In Computer Graphics, there is a strong research trail of methods for reproducing natural human motion by incorporating the true mechanics of the musculoskeletal system. However, common modeling assumptions such as the near-ubiquitous adoption of line-segment primitives for

Authors' addresses: Seunghwan Lee, Department of Computer Science and Engineering, Seoul National University, Seoul, [ls9021@mrl.snu.ac.kr](mailto:ls9021@mrl.snu.ac.kr); Ri Yu, Department of Computer Science and Engineering, Seoul National University, Seoul, [yuri@mrl.snu.ac.kr](mailto:yuri@mrl.snu.ac.kr); Jungnam Park, Department of Computer Science and Engineering, Seoul National University, Seoul, [jungnam04@mrl.snu.ac.kr](mailto:jungnam04@mrl.snu.ac.kr); Mridul Aanjaneya, Department of Computer Science, Rutgers University, New Brunswick, NJ, [mridul.aanjaneya@rutgers.edu](mailto:mridul.aanjaneya@rutgers.edu); Eftychios Sifakis, Computer Science Department, University of Wisconsin-Madison, Madison, WI, [sifakis@cs.wisc.edu](mailto:sifakis@cs.wisc.edu); Jehee Lee, Department of Computer Science and Engineering, Seoul National University, Seoul, [jehee@mrl.snu.ac.kr](mailto:jehee@mrl.snu.ac.kr).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.  
0730-0301/2018/8-ART57  
<https://doi.org/10.1145/3197517.3201330>

# The Human Touch: Measuring Contact with Real Human Soft Tissues

DINESH K. PAI, University of British Columbia and Vital Mechanics Research, Canada

AUSTIN ROTHWELL, PEARSON WYDER-HODGE, and ALISTAIR WICK, University of British Columbia

YE FAN and EGOR LARIONOV, University of British Columbia and Vital Mechanics Research

DARCY HARRISON, Vital Mechanics Research

DEBANGA RAJ NEOG and COLE SHING, University of British Columbia

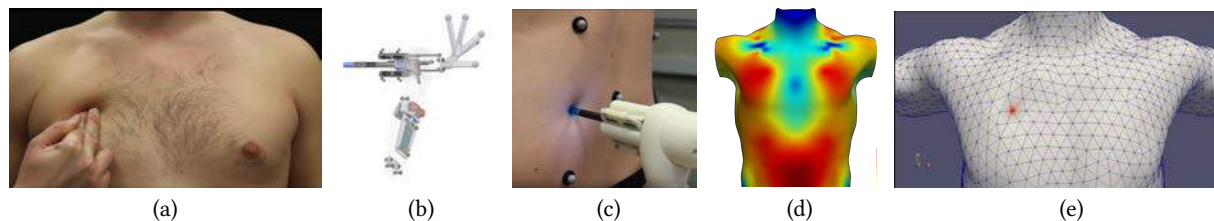


Fig. 1. The paper describes a complete pipeline for modeling contact with the human body (a). We developed a novel hand-held skin probe (b), incorporating a high resolution force/torque sensor and three cameras. The probe measures tissue properties at different locations on the body (c), while both probe and body are tracked by a Vicon motion capture system. The data are used to map the properties of our “sliding thick skin” tissue model over the participants body (d). Our data can be used to produce more realistic simulations of the human body in contact (e).

Simulating how the human body deforms in contact with external objects, tight clothing, or other humans is of central importance to many fields. Despite great advances in numerical methods, the material properties required to accurately simulate the body of a real human have been sorely lacking. Here we show that mechanical properties of the human body can be directly measured using a novel hand-held device. We describe a complete pipeline for measurement, modeling, parameter estimation, and simulation using the finite element method. We introduce a phenomenological model (the sliding thick skin model) that is effective for both simulation and parameter estimation. Our data also provide new insights into how the human body actually behaves. The methods described here can be used to create personalized models of an individual human or of a population. Consequently, our methods have many potential applications in computer animation, product design, e-commerce, and medicine.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: human simulation, contact, measurement, soft tissues, finite element, biomechanics

Authors' addresses: Dinesh K. Pai, pai@cs.ubc.ca, University of British Columbia, Computer Science, Vancouver, BC, Vital Mechanics Research, Vancouver, BC, Canada; Austin Rothwell, acroth@cs.ubc.ca; Pearson Wyder-Hodge, p.wyderhodge@alumni.ubc.ca; Alistair Wick, alistair.wk@gmail.com, University of British Columbia, Vancouver, BC; Ye Fan, yefan@vitalmechanics.com; Egor Larionov, egor.larionov@gmail.com, University of British Columbia, Computer Science, Vancouver, BC, Vital Mechanics Research, Vancouver, BC; Darcy Harrison, darcy.harrison@vitalmechanics.com, Vital Mechanics Research, Vancouver, BC; Debangaraj Neog, debanga@cs.ubc.ca; Cole Shing, coleshin@cs.ubc.ca, University of British Columbia, Computer Science, Vancouver, BC.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART58 \$15.00  
<https://doi.org/10.1145/3197517.3201296>

## ACM Reference Format:

Dinesh K. Pai, Austin Rothwell, Pearson Wyder-Hodge, Alistair Wick, Ye Fan, Egor Larionov, Darcy Harrison, Debangaraj Neog, and Cole Shing. 2018. The Human Touch: Measuring Contact with Real Human Soft Tissues. *ACM Trans. Graph.* 37, 4, Article 58 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201296>

## 1 INTRODUCTION

Simulating the human body is one of the greatest challenges in computer graphics, and also in myriad fields of human endeavor ranging from industrial design to medicine.

Most human models used in computer graphics are purely kinematic, and cannot simulate how the body would react to internal and external forces. There have been significant recent advances in physics-based models of the human body, but these have focused on the numerical methods required for simulating soft tissues; the constitutive properties of these tissues have been chosen almost arbitrarily. Even though such simulations are very useful for producing physically plausible jiggles and bulges, their utility outside the entertainment industry has been limited.

By contrast, applications in product design, e-commerce, and medicine need trustworthy models, whose predictions can be relied upon to make decisions. For example, before you buy a tight-fitting pair of jeans or a sports bra online, a trustworthy simulation should predict how well it fits you, taking into account how your body will deform when the garment is worn. Such accurate and personalized human body simulations are out of reach today.

Why is this? Despite intense research effort, the human body remains one of the most mysterious and complex objects known to science. Soft tissues are composites made of many different living materials whose properties can vary significantly over the body, across individuals, and even under voluntary control (e.g., by flexing a muscle). It is not currently possible to simply look up the mechanical properties of, say, your shoulder and plug it into a simulation.

# An Empirical Rig for Jaw Animation

GASPARD ZOSS, Disney Research

DEREK BRADLEY, Disney Research

PASCAL BÉRARD, Disney Research, ETH Zurich

THABO BEELER, Disney Research

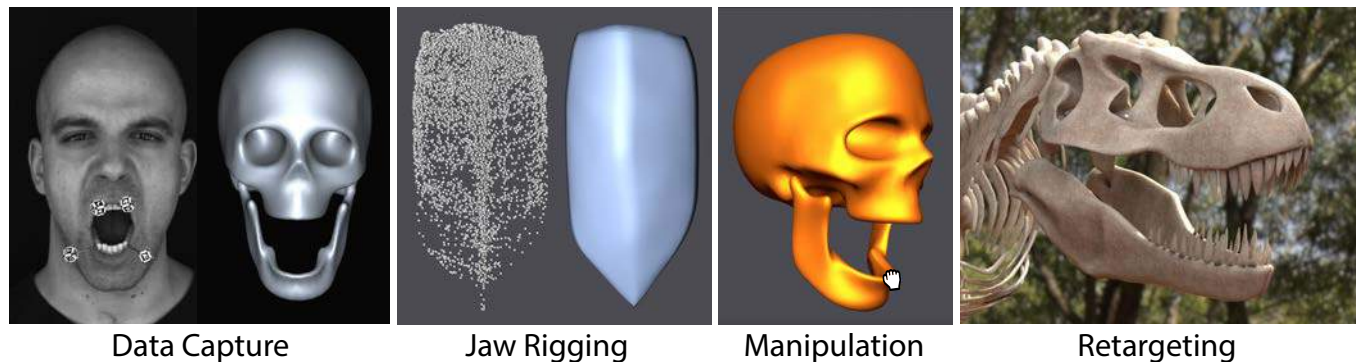


Fig. 1. We present an empirical rig for jaw animation, built from accurate capture data. Our rig is based on Posselt's Envelope of Motion, allowing intuitive 3-DOF control yet highly expressive and realistic jaw motions, and it can be retargeted to new characters and fantasy creatures.

In computer graphics the motion of the jaw is commonly modelled by up-down and left-right rotation around a fixed pivot plus a forward-backward translation, yielding a three dimensional rig that is highly suited for intuitive artistic control. The anatomical motion of the jaw is, however, much more complex since the joints that connect the jaw to the skull exhibit both rotational and translational components. In reality the jaw does not move in a three dimensional subspace but on a constrained manifold in six dimensions. We analyze this manifold in the context of computer animation and show how the manifold can be parameterized with three degrees of freedom, providing a novel jaw rig that preserves the intuitive control while providing more accurate jaw positioning. The chosen parameterization furthermore places anatomically correct limits on the motion, preventing the rig from entering physiologically infeasible poses. Our new jaw rig is empirically designed from accurate capture data, and we provide a simple method to retarget the rig to new characters, both human and fantasy.

CCS Concepts: • **Computing methodologies** → **Motion processing**; *Motion capture*;

Additional Key Words and Phrases: Jaw Rig, Data Driven Animation, Jaw Animation, Facial Animation, Motion Capture, Acquisition

Authors' addresses: Gaspard Zoss, Disney Research, [gaspard.zoss@disneyresearch.com](mailto:gaspard.zoss@disneyresearch.com); Derek Bradley, Disney Research, [derek.bradley@disneyresearch.com](mailto:derek.bradley@disneyresearch.com); Pascal Bérard, Disney Research, ETH Zurich, [pascal.berard@disneyresearch.com](mailto:pascal.berard@disneyresearch.com); Thabo Beeler, Disney Research, [thabo.beeler@disneyresearch.com](mailto:thabo.beeler@disneyresearch.com).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART59 \$15.00  
<https://doi.org/10.1145/3197517.3201382>

## ACM Reference Format:

Gaspard Zoss, Derek Bradley, Pascal Bérard, and Thabo Beeler. 2018. An Empirical Rig for Jaw Animation. *ACM Trans. Graph.* 37, 4, Article 59 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201382>

## 1 INTRODUCTION

When looking at the human face, the mandible (or *jaw-bone*) plays a central role in defining the facial structure and appearance. Its position fundamentally determines the shape of the skin as well as the placement of lower teeth, both of which are extremely important and salient visual features, and misplacement by even a few millimeters can be perceived. As a consequence, one of the most common orthognathic procedures is to extend or shorten the mandible by a few millimeters to correct for malocclusions, such as under- or over-bites. In computer graphics, the jaw plays a particularly important role during animation. Many facial rigs employ skinning to deform the skin as a function of the underlying bone motion, and hence it is important that this motion is correct. Such rigs employ abstract bones that are connected to each other via joints which limit their relative motion and offer an intuitive control structure.

The mandible is attached to the skull via the temporomandibular joint (TMJ), which is one of the most complex joints in the human body. Unlike a simple hinge joint (such as, for example, the elbow joint), the mandible slides over the surface of the skull bone while rotating, which means that the jaw does not have a fixed center of rotation (see Fig. 2). Furthermore, the final motion of the mandible is governed by the interplay of two such joints, one on each side of the head, with the consequence that the articulation takes place on a complex manifold in  $\mathbb{R}^6$ .

In computer animation this complexity is usually overlooked. Most commonly, animation rigs model the jaw joint by two rotations



# Tetrahedral Meshing in the Wild

YIXIN HU, New York University

QINGNAN ZHOU, Adobe Research

XIFENG GAO, New York University

ALEC JACOBSON, University of Toronto

DENIS ZORIN, New York University

DANIELE PANOZZO, New York University

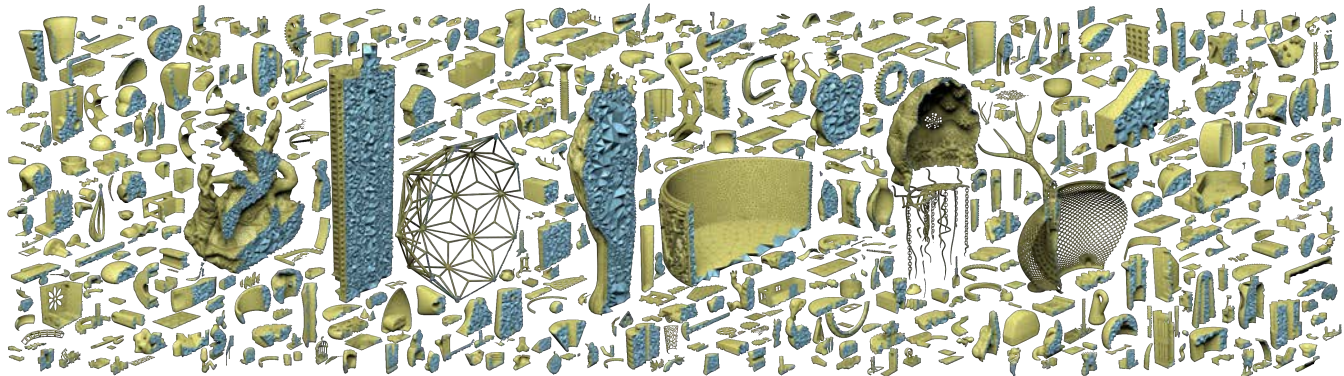


Fig. 1. A selection of the ten thousand meshes *in the wild* tetrahedralized by our novel tetrahedral meshing technique.

We propose a novel tetrahedral meshing technique that is unconditionally robust, requires no user interaction, and can directly convert a triangle soup into an analysis-ready volumetric mesh. The approach is based on several core principles: (1) initial mesh construction based on a fully robust, yet efficient, filtered exact computation (2) explicit (automatic or user-defined) tolerancing of the mesh relative to the surface input (3) iterative mesh improvement with guarantees, at every step, of the output validity. The quality of the resulting mesh is a direct function of the target mesh size and allowed tolerance: increasing allowed deviation from the initial mesh and decreasing the target edge length both lead to higher mesh quality.

Our approach enables “black-box” analysis, i.e. it allows to automatically solve partial differential equations on geometrical models available in the wild, offering a robustness and reliability comparable to, e.g., image processing algorithms, opening the door to automatic, large scale processing of real-world geometric data.

CCS Concepts: • **Mathematics of computing** → **Mesh generation**;

Additional Key Words and Phrases: Mesh Generation, Tetrahedral Meshing, Robust Geometry Processing

Authors' addresses: Yixin Hu, New York University, [yh1998@nyu.edu](mailto:yh1998@nyu.edu); Qingnan Zhou, Adobe Research, [qzhou@adobe.com](mailto:qzhou@adobe.com); Xifeng Gao, New York University, [gxf.xisha@gmail.com](mailto:gxf.xisha@gmail.com); Alec Jacobson, University of Toronto, [jacobson@cs.toronto.edu](mailto:jacobson@cs.toronto.edu); Denis Zorin, New York University, [dzorin@cs.nyu.edu](mailto:dzorin@cs.nyu.edu); Daniele Panozzo, New York University, [panozzo@nyu.edu](mailto:panozzo@nyu.edu).

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only.

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART60 \$15.00

<https://doi.org/10.1145/3197517.3201353>

## ACM Reference Format:

Yixin Hu, Qingnan Zhou, Xifeng Gao, Alec Jacobson, Denis Zorin, and Daniele Panozzo. 2018. Tetrahedral Meshing in the Wild. *ACM Trans. Graph.* 37, 4, Article 60 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201353>

## 1 INTRODUCTION

Triangulating the interior of a shape is a fundamental subroutine in 2D and 3D geometric computation.

For two-dimensional problems requiring meshing a domain, robust and efficient software for constrained Delaunay triangulation problem has been a tremendous boon to the development of robust and efficient automatic computational pipelines, in particular ones requiring solving PDEs. Robust 2D triangulations inside a given polygon boundary are also an essential spatial partitioning useful for fast point location, path traversal, and distance queries.

In 3D, the problem of robustly triangulating the interior of a given triangle surface mesh is just as well, if not more, motivated. While tremendous progress was made on various instances of the problem, it is far from solved by existing methods. While pipelines involving 3D *tetrahedralization* of smooth implicit surfaces are quite mature, pipelines using meshes as input either are limited to simple shapes or routinely fallback on manual intervention. The user may have to “fix” input surface meshes to cajole meshers to succeed due to unspoken pre-conditions, or output tetrahedral meshes must be repaired due to failure to meet basic post-conditions (such as manifoldness). Existing methods typically fail too often to support automatic pipelines, such as massive data processing for machine learning applications, or shape optimization. In many cases, while meshing may succeed, the size of the output mesh may be

# Curved Optimal Delaunay Triangulation

LEMAN FENG, Ecole des Ponts ParisTech  
PIERRE ALLIEZ, Université Côte d’Azur, Inria  
LAURENT BUSÉ, Université Côte d’Azur, Inria  
HERVÉ DELINGETTE, Université Côte d’Azur, Inria  
MATHIEU DESBRUN, Caltech

Mesheres with curvilinear elements hold the appealing promise of enhanced geometric flexibility and higher-order numerical accuracy compared to their commonly-used straight-edge counterparts. However, the generation of curved meshes remains a computationally expensive endeavor with current meshing approaches: high-order parametric elements are notoriously difficult to conform to a given boundary geometry, and enforcing a smooth and non-degenerate Jacobian everywhere brings additional numerical difficulties to the meshing of complex domains. In this paper, we propose an extension of Optimal Delaunay Triangulations (ODT) to curved and graded isotropic meshes. By exploiting a continuum mechanics interpretation of ODT instead of the usual approximation theoretical foundations, we formulate a very robust geometry and topology optimization of Bézier meshes based on a new simple functional promoting isotropic and uniform Jacobians throughout the domain. We demonstrate that our resulting curved meshes can adapt to complex domains with high precision even for a small count of elements thanks to the added flexibility afforded by more control points and higher order basis functions.

CCS Concepts: • Mathematics of computing → Mesh generation;

Additional Key Words and Phrases: Higher-order meshing, Optimal Delaunay Triangulations, higher order finite elements, Bézier elements.

## ACM Reference Format:

Leman Feng, Pierre Alliez, Laurent Busé, Hervé Delingette, and Mathieu Desbrun. 2018. Curved Optimal Delaunay Triangulation. *ACM Trans. Graph.* 37, 4, Article 61 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201358>

## 1 INTRODUCTION

Simplicial meshing has found wide adoption in the graphics community over the years because of the prevalence of linear basis functions in computer animation. As the availability of computing power on commodity hardware continues to increase, so is the geometric complexity of the models simulated. However, properly capturing the boundaries of complex shapes with linear elements often requires an inordinate count of simplices for highly curved domains. Computational efficiency, instead, dictates the use of a low count of higher order piecewise polynomial elements: while

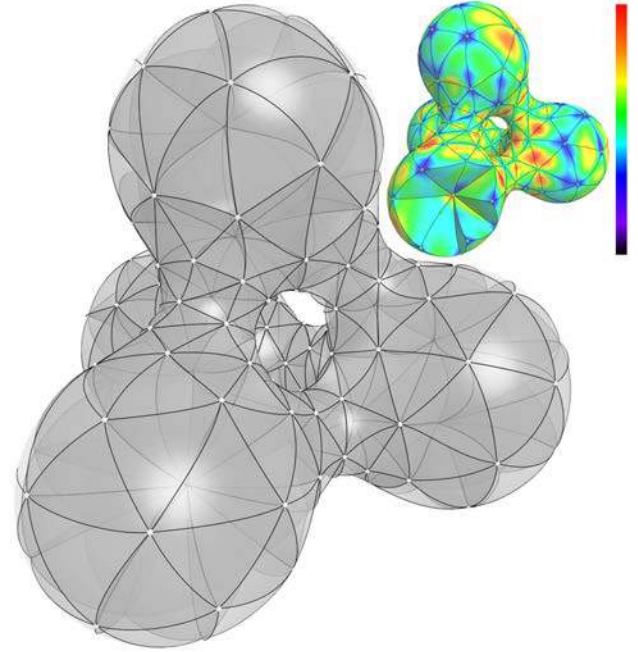


Fig. 1. **Curved Optimal Delaunay Triangulations:** A Bézier mesh (left, here with cubic patches) can capture a curved domain with orders of magnitude less elements for a given Hausdorff distance. Our extension of ODT to curved meshes results in a well-behaved Jacobian field (top: sizing-scaled determinant, with a few front elements removed to expose the interior).

using high-order elements—and thus, high-order basis functions—involves inevitable computational overhead, the drastic reduction of the number of elements needed to capture the domain boundary and the internal physical fields with a given accuracy lowers the total computational time appreciably compared to the linear case.

However, generating such a *curved* high order mesh for an arbitrary domain is significantly more difficult than its straight-edge analog: first, many staples of meshing (such as Delaunay triangulations or quality measures of mesh elements) only apply to straight-edge meshes; moreover, the use of higher-order polynomials to approximate the shape of a domain renders the numerical task more complex as it increases the occurrence of local minima in the energy landscape involved in the boundary fitting process and of local fold-overs of the polynomial map. As a result, most approaches for curved meshing limit themselves to deforming the boundary of a straight-edge mesh to improve the spatial matching of the boundary. However, this approach is limited and far from optimal as it largely relies on a good coarse linear mesh to start with, and leads to high deformation near the domain boundary, thus hampering subsequent numerical accuracy where it is sometimes most crucial (e.g., for boundary layers in fluid dynamics).

This paper extends the successful Optimal Delaunay Triangulation (ODT) approach for generating isotropic triangle and tetrahedron meshes [Alliez et al. 2005; Chen and Xu 2004] to now form Bézier meshes of arbitrary order. By exploiting the numerical properties of the original ODT technique and introducing a novel boundary

Authors’ addresses: L Feng, Ecole des Ponts ParisTech, France; P. Alliez, L. Busé, H. Delingette, Université Côte d’Azur / Inria, France; M. Desbrun, Caltech, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
0730-0301/2018/8-ART61 \$15.00  
<https://doi.org/10.1145/3197517.3201358>

# Computing a High-Dimensional Euclidean Embedding from an Arbitrary Smooth Riemannian Metric

ZICHUN ZHONG, Wayne State University

WENPING WANG, The University of Hong Kong

BRUNO LÉVY, INRIA Nancy - Grand Est

JING HUA, Wayne State University

XIAOHU GUO, University of Texas at Dallas

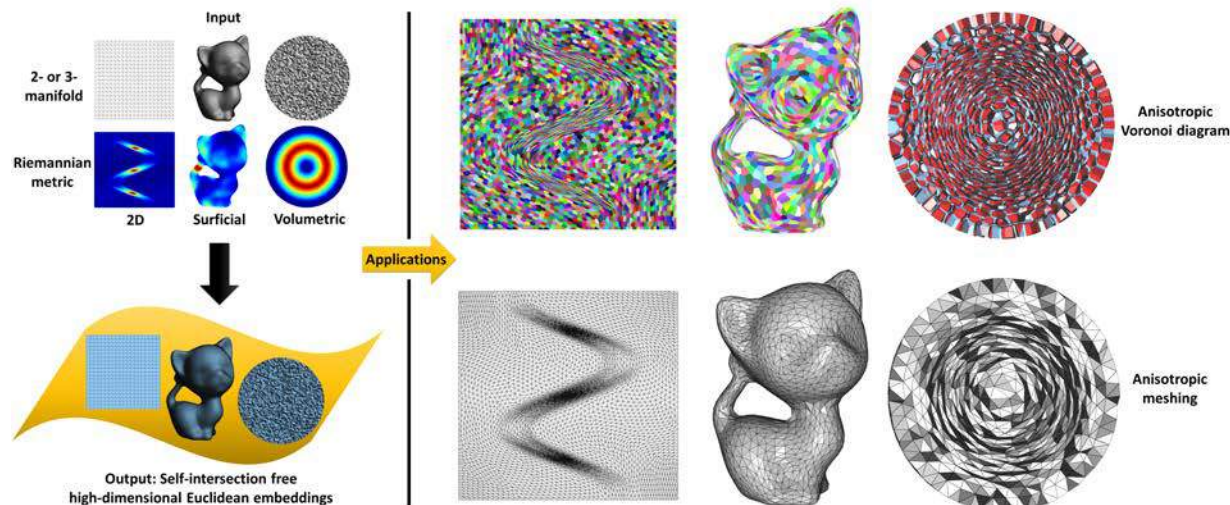


Fig. 1. Illustration of self-intersection free high-dimensional Euclidean embedding (SIFHDE<sup>2</sup>) for anisotropic Restricted Voronoi Diagram and meshing on Riemannian 2- and 3-manifolds. Left: For some examples of 2- and 3-manifolds with Riemannian metric fields in different scenarios, such as metrics defined by a 2D analytic function, 3D surface curvature, and 3D analytic function, we can compute SIFHDE<sup>2</sup>s of them in a high-dimensional Euclidean space. Right: Then, we can generate the Voronoi diagrams and meshes in such embedding with Euclidean metric. Finally, when they are mapped to the original domains, the anisotropic Voronoi diagrams (Top Right) and the anisotropic meshes (Bottom Right) will exhibit the desired anisotropy.

This article presents a new method to compute a *self-intersection free high-dimensional Euclidean embedding* (SIFHDE<sup>2</sup>) for surfaces and volumes equipped with an arbitrary Riemannian metric. It is already known that given a high-dimensional (high-d) embedding, one can easily compute an anisotropic Voronoi diagram by back-mapping it to 3D space. We show here how to solve the inverse problem, i.e., given an input metric, compute a smooth intersection-free high-d embedding of the input such that the pullback metric of the embedding matches the input metric. Our numerical solution mechanism matches the deformation gradient of the 3D  $\rightarrow$  higher-d mapping with the given Riemannian metric. We demonstrate the applicability of our

method, by using it to construct anisotropic Restricted Voronoi Diagram (RVD) and anisotropic meshing, that are otherwise extremely difficult to compute. In SIFHDE<sup>2</sup>-space constructed by our algorithm, difficult 3D anisotropic computations are replaced with simple Euclidean computations, resulting in an isotropic RVD and its dual mesh on this high-d embedding. Results are compared with the state-of-the-art in anisotropic surface and volume meshings using several examples and evaluation metrics.

**CCS Concepts:** • **Computing methodologies**  $\rightarrow$  **Parametric curve and surface models; Volumetric models;**

**Additional Key Words and Phrases:** anisotropic Voronoi diagram, anisotropic meshing, high-dimensional (high-d) embedding, self-intersection free

## ACM Reference Format:

Zichun Zhong, Wenping Wang, Bruno Lévy, Jing Hua, and Xiaohu Guo. 2018. Computing a High-Dimensional Euclidean Embedding from an Arbitrary Smooth Riemannian Metric. *ACM Trans. Graph.* 37, 4, Article 62 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201369>

## 1 INTRODUCTION

Anisotropic meshes are important for not only improving the accuracy of numerical simulations [Alauzet and Loseille 2010; Narain et al. 2012] but also better approximating shapes [Simpson 1994].

Authors' addresses: Zichun Zhong, Wayne State University, [zichunzhong@wayne.edu](mailto:zichunzhong@wayne.edu); Wenping Wang, The University of Hong Kong, [wenping@cs.hku.hk](mailto:wenping@cs.hku.hk); Bruno Lévy, INRIA Nancy - Grand Est, [bruno.levy@inria.fr](mailto:bruno.levy@inria.fr); Jing Hua, Wayne State University, [jinghua@wayne.edu](mailto:jinghua@wayne.edu); Xiaohu Guo, University of Texas at Dallas, [xguo@utdallas.edu](mailto:xguo@utdallas.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART62 \$15.00

<https://doi.org/10.1145/3197517.3201369>



# Shape from Metric

ALBERT CHERN, TU Berlin  
FELIX KNÖPPEL, TU Berlin  
ULRICH PINKALL, TU Berlin  
PETER SCHRÖDER, Caltech

We study the *isometric immersion* problem for orientable surface triangle meshes endowed with only a metric: given the combinatorics of the mesh together with edge lengths, approximate an isometric immersion into  $\mathbb{R}^3$ . To address this challenge we develop a discrete theory for surface immersions into  $\mathbb{R}^3$ . It precisely characterizes a discrete immersion, up to subdivision and small perturbations. In particular our discrete theory correctly represents the topology of the space of immersions, *i.e.*, the regular homotopy classes which represent its connected components. Our approach relies on unit quaternions to represent triangle orientations and to encode, in their parallel transport, the topology of the immersion. In unison with this theory we develop a computational apparatus based on a variational principle. Minimizing a non-linear Dirichlet energy optimally finds extrinsic geometry for the given intrinsic geometry and ensures low metric approximation error.

We demonstrate our algorithm with a number of applications from mathematical visualization and art directed isometric shape deformation, which mimics the behavior of thin materials with high membrane stiffness.

CCS Concepts: • **Mathematics of computing** → **Partial differential equations**; *Differential calculus*; *Algebraic topology*; • **Computing methodologies** → **Mesh models**; *Physical simulation*; • **Applied computing** → *Physics*;

Additional Key Words and Phrases: Differential geometry, geometry processing, isometric immersions, spin structures, quaternions.

## ACM Reference Format:

Albert Chern, Felix Knöppel, Ulrich Pinkall, and Peter Schröder. 2018. Shape from Metric. *ACM Trans. Graph.* 37, 4, Article 63 (August 2018), 17 pages. <https://doi.org/10.1145/3197517.3201276>

## 1 INTRODUCTION

Suppose for some abstract surface we only know its metric. Can you map it into  $\mathbb{R}^3$  in such a way that lengths of curves on the surface are preserved and small disks are realized in a 1-1 manner? We wish to solve this *isometric immersion* problem in the discrete setting when the surface is given in terms of a triangle mesh annotated with nothing more than lengths for every edge. A practical algorithm to numerically approximate such isometric immersions could have many applications. In Mathematical Visualization, we may encounter surfaces which are given purely in terms of their

Authors' addresses: Albert Chern, TU Berlin; Felix Knöppel, TU Berlin; Ulrich Pinkall, TU Berlin; Peter Schröder, Caltech.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART63 \$15.00

<https://doi.org/10.1145/3197517.3201276>

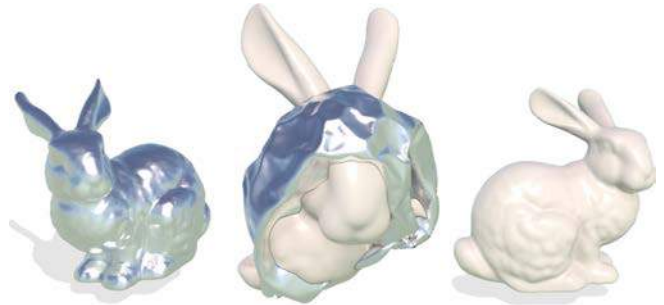


Fig. 1. Turning the bunny inside out while preserving its metric. See the video at 02:32.

metric without any explicit immersion (Fig. 18). In Geometry Processing we may be interested in isometric deformations which stay immersed (Fig. 1). Or, if the metric is manipulated, be it in shape space interpolation, editing, metric flows (Fig. 15), or others, we may need to find an immersion realizing the new metric. And in art directed form finding the characteristic “crumpling” behaviour of close-to-isometric surfaces can be of value (Figs. 12 & 13).

To pursue a program of numerically approximating isometric immersions we of course need to know that they exist in the first place. Fortunately this is the case both in the  $C^1$  (once continuously differentiable) setting [Kuiper 1955; Nash 1954] and in the PL (piecewise linear) setting [Burago and Zalgaller 1960, 1995] (the PL setting also allows for additional piecewise linear subdivisions).

While these theorems ensure that isometric immersions always exist and are even deformable, the respective constructions also imply that there is an isometry arbitrarily near *any* (possibly scaled) immersion. Among this multitude of possibilities (Fig. 2) we want to

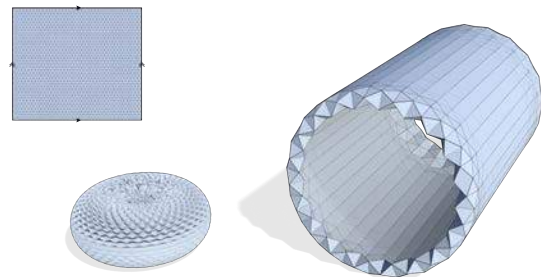


Fig. 2. Two isometric immersions of a flat torus (upper left, not to scale) using extensive “crumpling” (lower left) and nearly none (right). Both immersions of the torus are approximately isometric yet the “crumpled” torus has a very different appearance and apparent size.

# Synthetic Depth-of-Field with a Single-Camera Mobile Phone

NEAL WADHWA, RAHUL GARG, DAVID E. JACOBS, BRYAN E. FELDMAN, NORI KANAZAWA, ROBERT CARROLL, YAIR MOVSHOVITZ-ATTIAS, JONATHAN T. BARRON, Yael PRITCH, and MARC LEVOY, Google Research

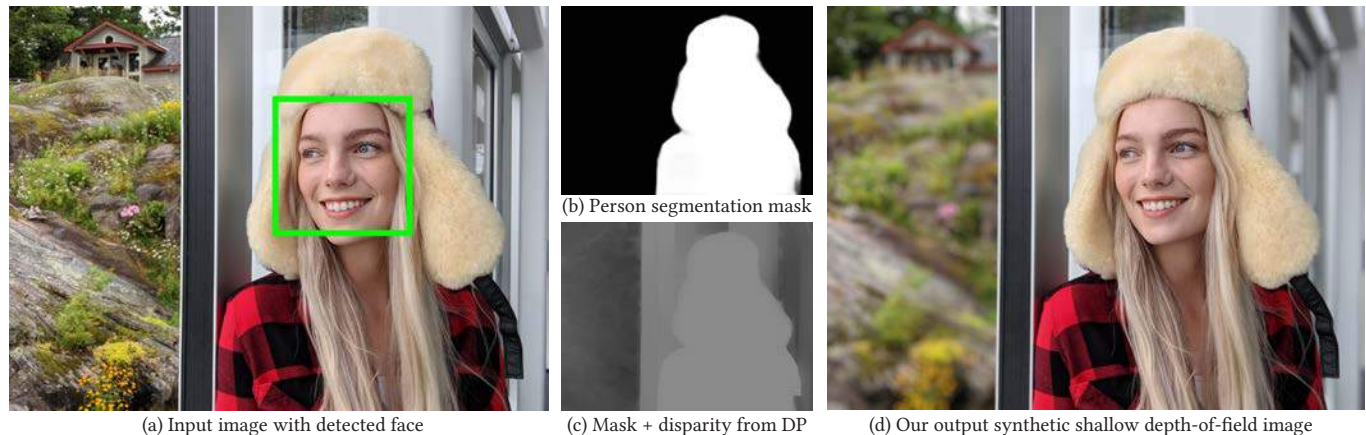


Fig. 1. We present a system that uses a person segmentation mask (b) and a noisy depth map computed using the camera's dual-pixel (DP) auto-focus hardware (c) to produce a synthetic shallow depth-of-field image (d) with a depth-dependent blur on a mobile phone. Our system is marketed as "Portrait Mode" on several Google-branded phones.

Shallow depth-of-field is commonly used by photographers to isolate a subject from a distracting background. However, standard cell phone cameras cannot produce such images optically, as their short focal lengths and small apertures capture nearly all-in-focus images. We present a system to computationally synthesize shallow depth-of-field images with a single mobile camera and a single button press. If the image is of a person, we use a person segmentation network to separate the person and their accessories from the background. If available, we also use dense *dual-pixel* auto-focus hardware, effectively a 2-sample light field with an approximately 1 millimeter baseline, to compute a dense depth map. These two signals are combined and used to render a defocused image. Our system can process a 5.4 megapixel image in 4 seconds on a mobile phone, is fully automatic, and is robust enough to be used by non-experts. The modular nature of our system allows it to degrade naturally in the absence of a dual-pixel sensor or a human subject.

CCS Concepts: • **Computing methodologies** → **Computational photography**; *Image processing*;

Additional Key Words and Phrases: depth-of-field, defocus, stereo, segmentation

Authors' address: Neal Wadhwa; Rahul Garg; David E. Jacobs; Bryan E. Feldman; Nori Kanazawa; Robert Carroll; Yair Movshovitz-Attias; Jonathan T. Barron; Yael Pritch; Marc Levoy Google Research, 1600 Amphitheater Parkway, Mountain View, CA, 94043.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART64

<https://doi.org/10.1145/3197517.3201329>

## ACM Reference Format:

Neal Wadhwa, Rahul Garg, David E. Jacobs, Bryan E. Feldman, Nori Kanazawa, Robert Carroll, Yair Movshovitz-Attias, Jonathan T. Barron, Yael Pritch, and Marc Levoy. 2018. Synthetic Depth-of-Field with a Single-Camera Mobile Phone. *ACM Trans. Graph.* 37, 4, Article 64 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201329>

## 1 INTRODUCTION

Depth-of-field is an important aesthetic quality of photographs. It refers to the range of depths in a scene that are imaged sharply in focus. This range is determined primarily by the aperture of the capturing camera's lens: a wide aperture produces a shallow (small) depth-of-field, while a narrow aperture produces a wide (large) depth-of-field. Professional photographers frequently use depth-of-field as a compositional tool. In portraiture, for instance, a strong background blur and shallow depth-of-field allows the photographer to isolate a subject from a cluttered, distracting background. The hardware used by DSLR-style cameras to accomplish this effect also makes these cameras expensive, inconvenient, and often difficult to use. Therefore, the compelling images they produce are largely limited to professionals. Mobile phone cameras are ubiquitous, but their lenses have apertures too small to produce the same kinds of images optically.

Recently, mobile phone manufacturers have started computationally producing shallow depth-of-field images. The most common technique is to include two cameras instead of one and to apply stereo algorithms to captured image pairs to compute a depth map. One of the images is then blurred according to this depthmap. However, adding a second camera raises manufacturing costs, increases

# Stereo Magnification: Learning view synthesis using multiplane images

TINGHUI ZHOU, University of California, Berkeley

RICHARD TUCKER, Google

JOHN FLYNN, Google

GRAHAM FYFFE, Google

NOAH SNAVELY, Google

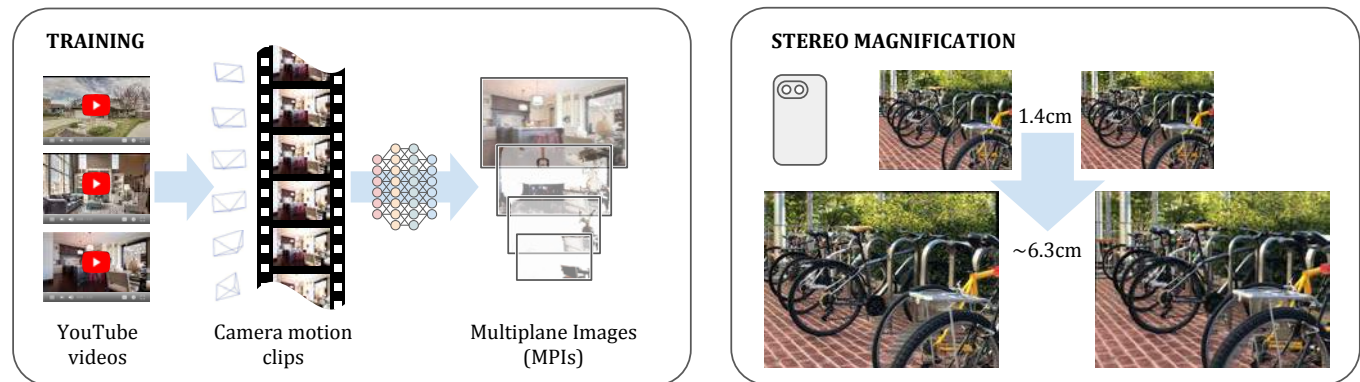


Fig. 1. We extract camera motion clips from YouTube videos and use them to train a neural network to generate a Multiplane Image (MPI) scene representation from narrow-baseline stereo image pairs. The inferred MPI representation can then be used to synthesize novel views of the scene, including ones that extrapolate significantly beyond the input baseline. (Video stills in this and other figures are used under Creative-Commons license from YouTube user *SonaVisual*.)

The view synthesis problem—generating novel views of a scene from known imagery—has garnered recent attention due in part to compelling applications in virtual and augmented reality. In this paper, we explore an intriguing scenario for view synthesis: extrapolating views from imagery captured by narrow-baseline stereo cameras, including VR cameras and now-widespread dual-lens camera phones. We call this problem *stereo magnification*, and propose a learning framework that leverages a new layered representation that we call *multiplane images* (MPIs). Our method also uses a massive new data source for learning view extrapolation: online videos on YouTube. Using data mined from such videos, we train a deep network that predicts an MPI from an input stereo image pair. This inferred MPI can then be used to synthesize a range of novel views of the scene, including views that extrapolate significantly beyond the input baseline. We show that our method compares favorably with several recent view synthesis methods, and demonstrate applications in magnifying narrow-baseline stereo images.

CCS Concepts: • **Computing methodologies** → **Computational photography**; **Image-based rendering**; **Neural networks**; **Virtual reality**;

Additional Key Words and Phrases: View extrapolation, deep learning

Authors' addresses: Tinghui Zhou, University of California, Berkeley; Richard Tucker, Google; John Flynn, Google; Graham Fyffe, Google; Noah Snavely, Google.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART65

<https://doi.org/10.1145/3197517.3201323>

## ACM Reference Format:

Tinghui Zhou, Richard Tucker, John Flynn, Graham Fyffe, and Noah Snavely. 2018. Stereo Magnification: Learning view synthesis using multiplane images. *ACM Trans. Graph.* 37, 4, Article 65 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201323>

## 1 INTRODUCTION

Photography has undergone an upheaval over the past decade. Cell-phone cameras have steadily displaced point-and-shoot cameras, and have become competitive with digital SLRs in certain scenarios. This change has been driven by the increasing image quality of cellphone cameras, through better hardware and also through computational photography functionality such as high dynamic range imaging [Hasinoff et al. 2016] and synthetic defocus [Apple 2016; Google 2017b]. Many of these recent innovations have sought to replicate capabilities of traditional cameras. However, cell phones are also rapidly acquiring new kinds of sensors, such as multiple lenses and depth sensors, enabling applications beyond traditional photography.

In particular, dual-lens cameras are becoming increasingly common. While stereo cameras have been around for nearly as long as photography itself, recently a number of dual-camera phones, such as the iPhone 7, have appeared on the market. These cameras tend to have a very small baseline (distance between views) on the order of a centimeter. We have also seen the recent appearance of a number of “virtual-reality ready” cameras that capture stereo images and



# In the Blink of an Eye – Leveraging Blink-Induced Suppression for Imperceptible Position and Orientation Redirection in Virtual Reality

EIKE LANGBEHN, University of Hamburg, Germany  
FRANK STEINICKE, University of Hamburg, Germany  
MARKUS LAPPE, University of Muenster, Germany  
GREGORY F. WELCH, University of Central Florida, USA  
GERD BRUDER, University of Central Florida, USA

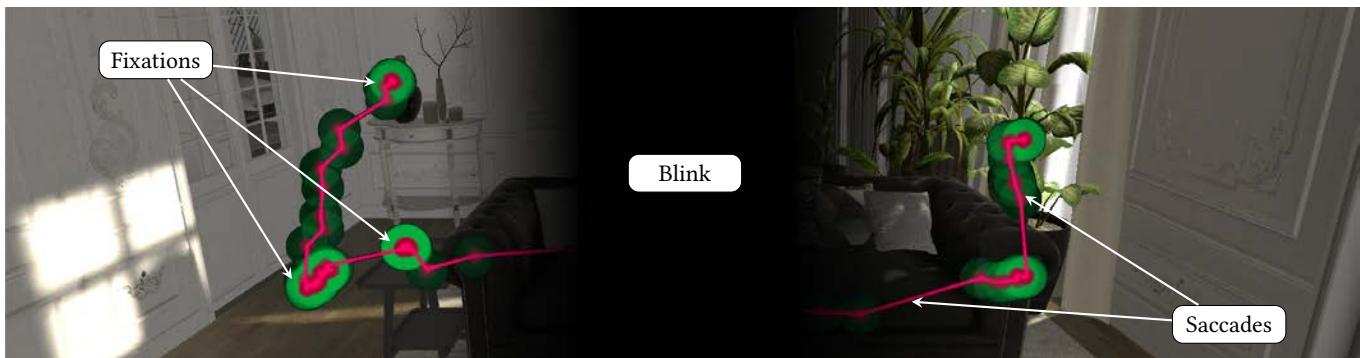


Fig. 1. Illustration of the viewing behavior of a virtual reality user including fixations (green) and saccades (red). A blink fully suppresses visual perception.

Immersive computer-generated environments (aka *virtual reality*, VR) are limited by the physical space around them, e. g., enabling natural walking in VR is only possible by perceptually-inspired locomotion techniques such as redirected walking (RDW). We introduce a completely new approach to imperceptible position and orientation redirection that takes advantage of the fact that even healthy humans are functionally blind for circa ten percent of the time under normal circumstances due to motor processes preventing light from reaching the retina (such as eye blinks) or perceptual processes suppressing degraded visual information (such as blink-induced suppression). During such periods of missing visual input, *change blindness* occurs, which denotes the inability to perceive a visual change such as the motion of an object or self-motion of the observer. We show that this phenomenon can be exploited in VR by synchronizing the computer graphics rendering system with the human visual processes for imperceptible camera movements, in particular to implement position and orientation redirection. We analyzed

human sensitivity to such visual changes with detection thresholds, which revealed that commercial off-the-shelf eye trackers and head-mounted displays suffice to translate a user by circa 4 - 9 cm and rotate the user by circa 2 - 5 degrees in any direction, which could be accumulated each time the user blinks. Moreover, we show the potential for RDW, whose performance could be improved by approximately 50 % when using our technique.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; • **Computing methodologies** → *Perception*; *Virtual reality*;

Additional Key Words and Phrases: Virtual reality, eye blinks, redirected walking, psychophysical experiments

## ACM Reference Format:

Eike Langbehn, Frank Steinicke, Markus Lappe, Gregory F. Welch, and Gerd Bruder. 2018. In the Blink of an Eye – Leveraging Blink-Induced Suppression for Imperceptible Position and Orientation Redirection in Virtual Reality. *ACM Trans. Graph.* 37, 4, Article 66 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201335>

## 1 INTRODUCTION

Locomotion, the act of moving from one location to another, is the driving problem of our research as it is considered one of the most fundamental and important activities performed during interaction with our surroundings, and remains a difficult challenge to be solved in the research field of virtual reality (VR), in which users are immersed in a computer-generated virtual environment (VE). Together with continuous improvements of VR technologies over the last decades, we have seen the shift from artificial to more natural locomotion interfaces in this context [Steinicke et al. 2013]. In

Authors' addresses: Eike Langbehn, University of Hamburg, Human-Computer Interaction, Vogt-Kölln-Str. 30, Hamburg, Hamburg, 22761, Germany; Frank Steinicke, University of Hamburg, Human-Computer Interaction, Vogt-Kölln-Str. 30, Hamburg, Hamburg, 22761, Germany; Markus Lappe, University of Muenster, Department of Psychology, Muenster, Germany; Gregory F. Welch, University of Central Florida, Synthetic Reality Lab, Orlando, USA; Gerd Bruder, University of Central Florida, Synthetic Reality Lab, Orlando, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
0730-0301/2018/8-ART66 \$15.00  
<https://doi.org/10.1145/3197517.3201335>

# Towards Virtual Reality Infinite Walking: Dynamic Saccadic Redirection

QI SUN, Stony Brook University, NVIDIA, and Adobe Research

ANJUL PATNEY, NVIDIA

LI-YI WEI, Adobe Research

OMER SHAPIRA, NVIDIA

JINGWAN LU and PAUL ASENTE, Adobe Research

SUWEN ZHU, Stony Brook University

MORGAN MCGUIRE and DAVID LUEBKE, NVIDIA

ARIE KAUFMAN, Stony Brook University

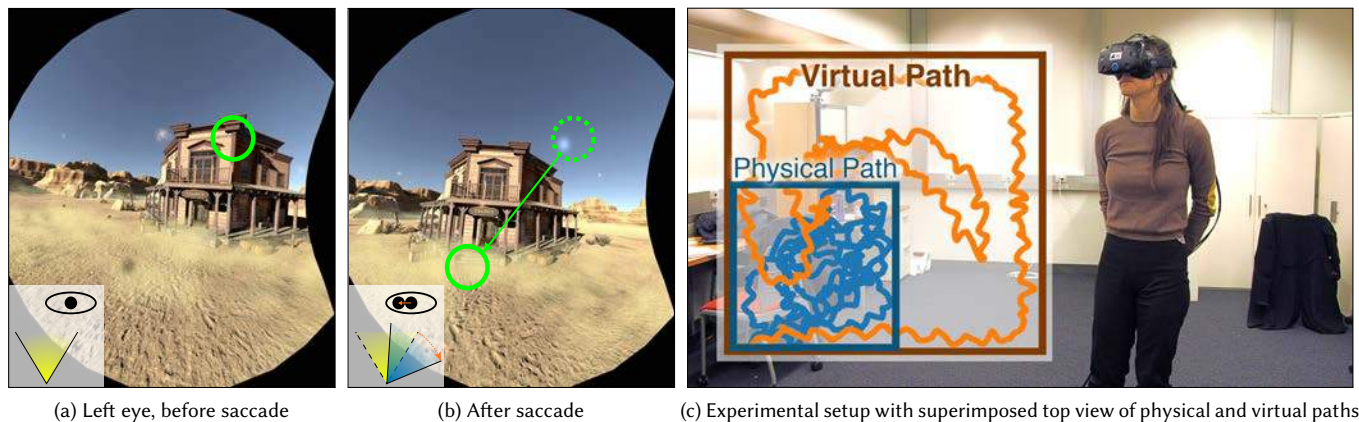


Fig. 1. *Triggering and harnessing temporary blindness via saccades for room-scale redirected walking in VR.* Our system renders a virtual environment into a pair of HMD views while tracking the user eye gaze. (a) shows a HMD left eye rendering for the viewer with overlaid visualizations of tracked eye gaze (green circle) and view frustum (lower left corner). When saccades (rapid eye movements) and head rotations are detected, our system rotates the virtual environments to redirect the users (b). Such rotations are visible during normal viewing conditions, but can be imperceptible during eye or head movements. (c) photographs our experimental setup with a Vive HMD augmented with SMI gaze tracking. Superimposed are the top view of the recorded movements of the physical path in a  $3.5\text{m} \times 3.5\text{m}$  real room and the virtual path in a much larger  $6.4\text{m} \times 6.4\text{m}$  synthetic space. Scene courtesy of NOT\_Lonely (Vitaly).

Redirected walking techniques can enhance the immersion and visual-vestibular comfort of virtual reality (VR) navigation, but are often limited by the size, shape, and content of the physical environments.

We propose a redirected walking technique that can apply to small physical environments with static or dynamic obstacles. Via a head- and eye-tracking VR headset, our method detects saccadic suppression and redirects the users during the resulting temporary blindness. Our dynamic path planning runs in real-time on a GPU, and thus can avoid static and dynamic obstacles, including walls, furniture, and other VR users sharing the same

physical space. To further enhance saccadic redirection, we propose subtle gaze direction methods tailored for VR perception.

We demonstrate that saccades can significantly increase the rotation gains during redirection without introducing visual distortions or simulator sickness. This allows our method to apply to large open virtual spaces and small physical environments for room-scale VR. We evaluate our system via numerical simulations and real user studies.

CCS Concepts: • **Human-centered computing** → **Virtual reality**; • **Computing methodologies** → **Perception**;

Additional Key Words and Phrases: virtual reality, redirected walking, human perception, saccade

## ACM Reference Format:

Qi Sun, Anjul Patney, Li-Yi Wei, Omer Shapira, Jingwan Lu, Paul Asente, Suwen Zhu, Morgan McGuire, David Luebke, and Arie Kaufman. 2018. Towards Virtual Reality Infinite Walking: Dynamic Saccadic Redirection. *ACM Trans. Graph.* 37, 4, Article 67 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201294>

Authors' addresses: Qi Sun, Stony Brook University, NVIDIA, Adobe Research; Anjul Patney, NVIDIA; Li-Yi Wei, Adobe Research; Omer Shapira, NVIDIA; Jingwan Lu; Paul Asente, Adobe Research; Suwen Zhu, Stony Brook University; Morgan McGuire; David Luebke, NVIDIA; Arie Kaufman, Stony Brook University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART67 \$15.00 <https://doi.org/10.1145/3197517.3201294>

# Deep Appearance Models for Face Rendering

STEPHEN LOMBARDI, Facebook Reality Labs

JASON SARAGIH, Facebook Reality Labs

TOMAS SIMON, Facebook Reality Labs

YASER SHEIKH, Facebook Reality Labs

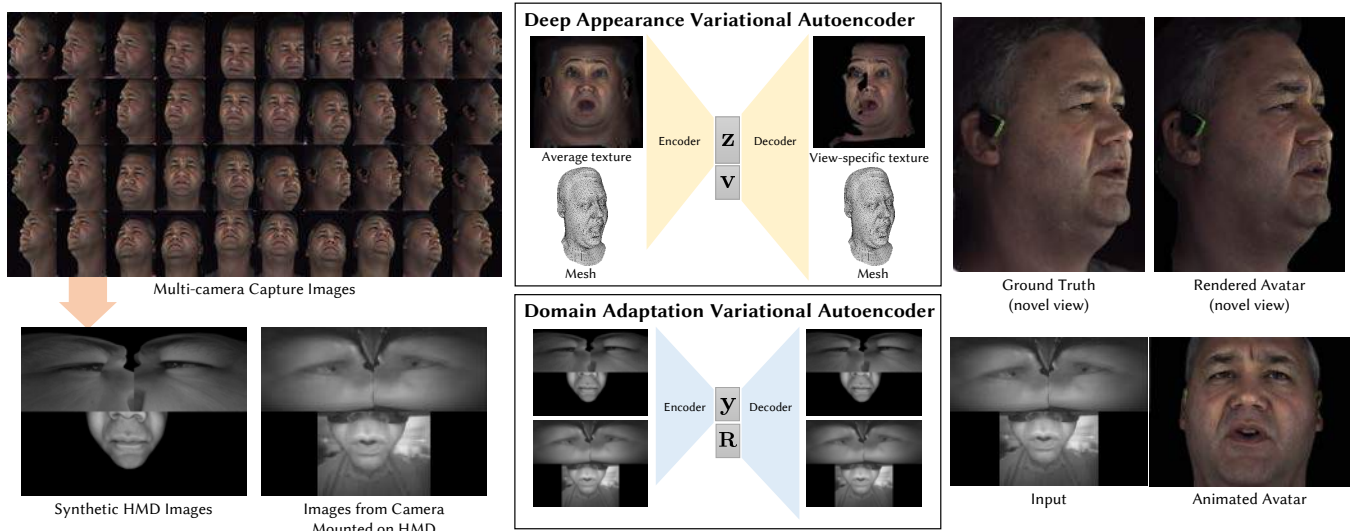


Fig. 1. Our model jointly encodes and decodes geometry and view-dependent appearance into a latent code  $z$ , from data captured from a multi-camera rig, enabling highly realistic data-driven facial rendering. We use this rich data to drive our avatars from cameras mounted on a head-mounted display (HMD). We do this by creating synthetic HMD images through image-based rendering, and using another variational autoencoder to learn a common representation  $y$  of real and synthetic HMD images. We then regress from  $y$  to the latent rendering code  $z$  and decode into mesh and texture to render. Our method enables high-fidelity social interaction in virtual reality.

We introduce a deep appearance model for rendering the human face. Inspired by Active Appearance Models, we develop a data-driven rendering pipeline that learns a joint representation of facial geometry and appearance from a multiview capture setup. Vertex positions and view-specific textures are modeled using a deep variational autoencoder that captures complex non-linear effects while producing a smooth and compact latent representation. View-specific texture enables the modeling of view-dependent effects such as specularities. In addition, it can also correct for imperfect geometry stemming from biased or low resolution estimates. This is a significant departure from the traditional graphics pipeline, which requires highly accurate geometry as well as all elements of the shading model to achieve realism through physically-inspired light transport. Acquiring such a high level of accuracy is difficult in practice, especially for complex and intricate parts of the face,

Authors' addresses: Stephen Lombardi, Facebook Reality Labs, Pittsburgh, PA, [stephen.lombardi@fb.com](mailto:stephen.lombardi@fb.com); Jason Saragih, Facebook Reality Labs, Pittsburgh, PA, [jason.saragih@fb.com](mailto:jason.saragih@fb.com); Tomas Simon, Facebook Reality Labs, Pittsburgh, PA, [tomas.simon@fb.com](mailto:tomas.simon@fb.com); Yaser Sheikh, Facebook Reality Labs, Pittsburgh, PA, [yaser.sheikh@fb.com](mailto:yaser.sheikh@fb.com).



This work is licensed under a Creative Commons Attribution International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART68  
<https://doi.org/10.1145/3197517.3201401>

such as eyelashes and the oral cavity. These are handled naturally by our approach, which does not rely on precise estimates of geometry. Instead, the shading model accommodates deficiencies in geometry through the flexibility afforded by the neural network employed. At inference time, we condition the decoding network on the viewpoint of the camera in order to generate the appropriate texture for rendering. The resulting system can be implemented simply using existing rendering engines through dynamic textures with flat lighting. This representation, together with a novel unsupervised technique for mapping images to facial states, results in a system that is naturally suited to real-time interactive settings such as Virtual Reality (VR).

CCS Concepts: • **Computing methodologies** → **Image-based rendering**; *Neural networks*; *Virtual reality*; *Mesh models*;

Additional Key Words and Phrases: Appearance Models, Image-based Rendering, Deep Appearance Models, Face Rendering

## ACM Reference Format:

Stephen Lombardi, Jason Saragih, Tomas Simon, and Yaser Sheikh. 2018. Deep Appearance Models for Face Rendering. *ACM Trans. Graph.* 37, 4, Article 68 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201401>

## 1 INTRODUCTION

With the advent of modern Virtual Reality (VR) headsets, there is a need for improved real-time computer graphics models for enhanced immersion. Traditional computer graphics techniques are



# Neural Best-Buddies: Sparse Cross-Domain Correspondence

KFIR ABERMAN, AICFVE Beijing Film Academy, Tel-Aviv University

JING LIAO, Microsoft Research Asia

MINGYI SHI, Shandong University

DANI LISCHINSKI, Hebrew University of Jerusalem

BAOQUAN CHEN, Shandong University, AICFVE Beijing Film Academy, Peking University

DANIEL COHEN-OR, Tel-Aviv University

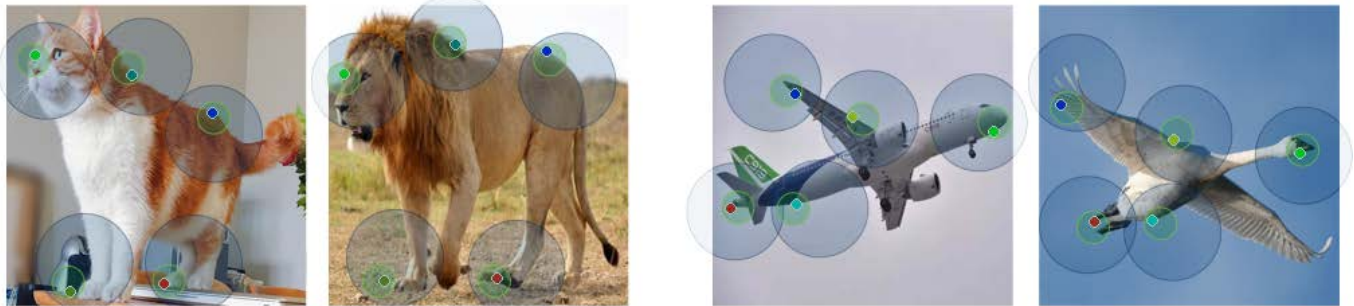


Fig. 1. Top 5 Neural Best-Buddies for two cross-domain image pairs. Using deep features of a pre-trained neural network, our coarse-to-fine sparse correspondence algorithm first finds high-level, low resolution, semantically matching areas (indicated by the large blue circles), then narrows down the search area to intermediate levels (middle green circles), until precise localization on well-defined edges in the pixel space (colored in corresponding unique colors). Image credits: PASCAL 3D+ dataset (second right image), DanZoo (second left and rightmost images).

Correspondence between images is a fundamental problem in computer vision, with a variety of graphics applications. This paper presents a novel method for *sparse cross-domain correspondence*. Our method is designed for pairs of images where the main objects of interest may belong to different semantic categories and differ drastically in shape and appearance, yet still contain semantically related or geometrically similar parts. Our approach operates on hierarchies of deep features, extracted from the input images by a pre-trained CNN. Specifically, starting from the coarsest layer in both hierarchies, we search for Neural Best Buddies (NBB): pairs of neurons that are mutual nearest neighbors. The key idea is then to percolate NBBs through the hierarchy, while narrowing down the search regions at each level and retaining only NBBs with significant activations. Furthermore, in order to overcome differences in appearance, each pair of search regions is transformed into a common appearance.

We evaluate our method via a user study, in addition to comparisons with alternative correspondence approaches. The usefulness of our method is demonstrated using a variety of graphics applications, including cross-domain image alignment, creation of hybrid images, automatic image morphing, and more.

CCS Concepts: • **Computing methodologies** → *Interest point and salient region detections; Matching; Image manipulation;*

Authors' addresses: Kfir Aberman, AICFVE Beijing Film Academy, Tel-Aviv University; Jing Liao, Microsoft Research Asia; Mingyi Shi, Shandong University; Dani Lischinski, Hebrew University of Jerusalem; Baoquan Chen, Shandong University, AICFVE Beijing Film Academy, Peking University; Daniel Cohen-Or, Tel-Aviv University.

© 2018 Association for Computing Machinery.

This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in *ACM Transactions on Graphics*, <https://doi.org/10.1145/3197517.3201332>.

Additional Key Words and Phrases: cross-domain correspondence, image hybrids, image morphing

## ACM Reference Format:

Kfir Aberman, Jing Liao, Mingyi Shi, Dani Lischinski, Baoquan Chen, and Daniel Cohen-Or. 2018. Neural Best-Buddies: Sparse Cross-Domain Correspondence. *ACM Trans. Graph.* 37, 4, Article 69 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201332>

## 1 INTRODUCTION

Finding correspondences between a pair of images has been a long standing problem, with a multitude of applications in computer vision and graphics. In particular, sparse sets of corresponding point pairs may be used for tasks such as template matching, image alignment, and image morphing, to name a few. Over the years, a variety of dense and sparse correspondence methods have been developed, most of which assume that the input images depict the same scene or object (with differences in viewpoint, lighting, object pose, etc.), or a pair of objects from the same class.

In this work, we are concerned with *sparse cross-domain correspondence*: a more general and challenging version of the sparse correspondence problem, where the object of interest in the two input images can differ more drastically in their shape and appearance, such as objects belonging to different semantic categories (domains). It is, however, assumed that the objects contain at least some semantically related parts or geometrically similar regions, otherwise the correspondence task cannot be considered well-defined. Two examples of cross-domain scenarios and the results of our approach are shown in Figure 1. We focus on sparse correspondence,

# Deep Convolutional Priors for Indoor Scene Synthesis

KAI WANG, Brown University  
MANOLIS SAVVA, Princeton University  
ANGEL X. CHANG, Princeton University  
DANIEL RITCHIE, Brown University

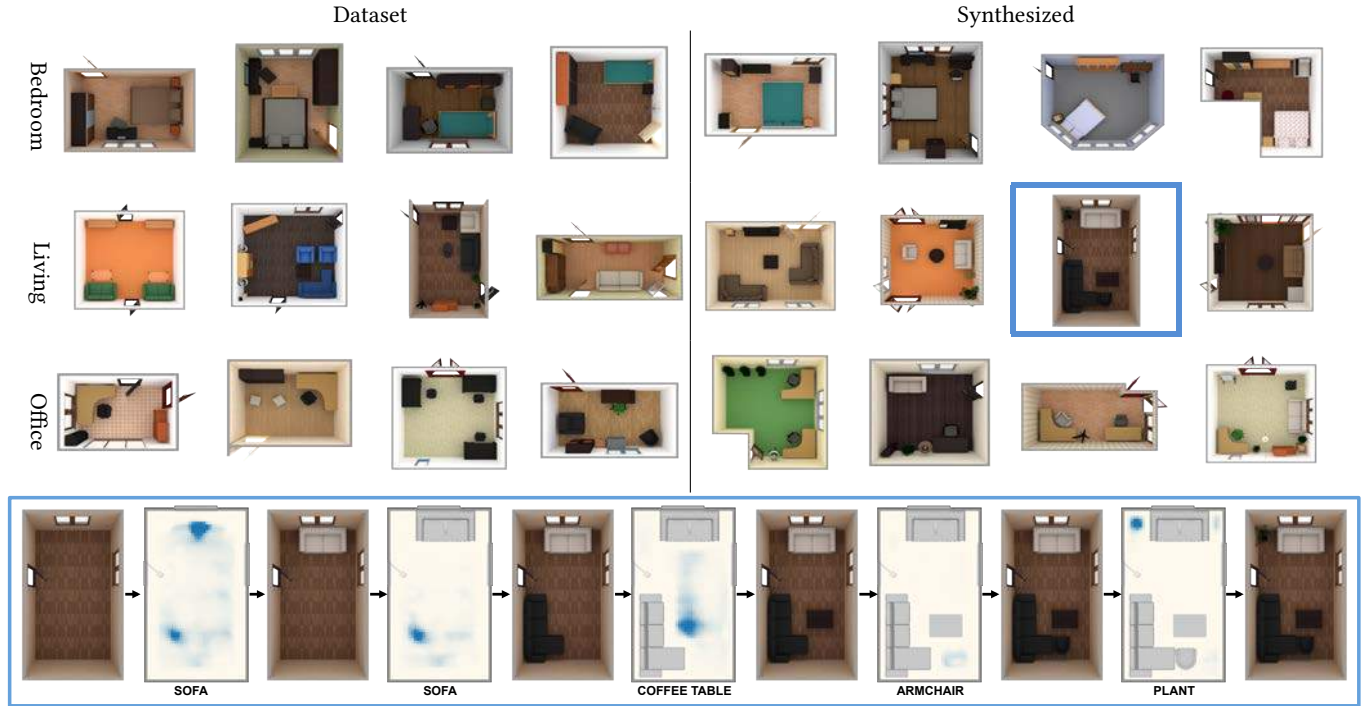


Fig. 1. We present a method to synthesize indoor scenes using deep convolutional network priors trained on top-down views of scenes from a scene dataset. *Top*: Bedroom, living room, and office scenes from our dataset, along with scenes synthesized from our model. *Bottom*: Our model generates scenes by iteratively inserting one object at a time. Here we show one such object insertion sequence for the living room scene outlined in blue. Renders of intermediate scenes are interleaved with a visualization of our model's predicted spatial probability distribution for the type of object about to be inserted.

We present a convolutional neural network based approach for indoor scene synthesis. By representing 3D scenes with a semantically-enriched image-based representation based on orthographic top-down views, we learn convolutional object placement priors from the entire context of a room. Our approach iteratively generates rooms from scratch, given only the room architecture as input. Through a series of perceptual studies we compare the plausibility of scenes generated using our method against baselines for

object selection and object arrangement, as well as scenes modeled by people. We find that our method generates scenes that are preferred over the baselines, and in some cases are equally preferred to human-created scenes.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Neural networks**; **Probabilistic reasoning**;

Additional Key Words and Phrases: scene synthesis, object layout, neural networks, convolutional networks, deep learning

## ACM Reference Format:

Kai Wang, Manolis Savva, Angel X. Chang, and Daniel Ritchie. 2018. Deep Convolutional Priors for Indoor Scene Synthesis. *ACM Trans. Graph.* 37, 4, Article 70 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201362>

## 1 INTRODUCTION

People spend a large percentage of their lives indoors: in offices, living rooms, bedrooms, and other interior spaces. A number of industries use virtual reproductions of such indoor scenes: interior design, architecture, gaming and virtual reality are a few. A

Authors' addresses: Kai Wang, Brown University; Manolis Savva, Princeton University; Angel X. Chang, Princeton University; Daniel Ritchie, Brown University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART70 \$15.00

<https://doi.org/10.1145/3197517.3201362>

# Point Convolutional Neural Networks by Extension Operators

MATAN ATZMON\*, Weizmann Institute of Science  
HAGGAI MARON\*, Weizmann Institute of Science  
YARON LIPMAN, Weizmann Institute of Science

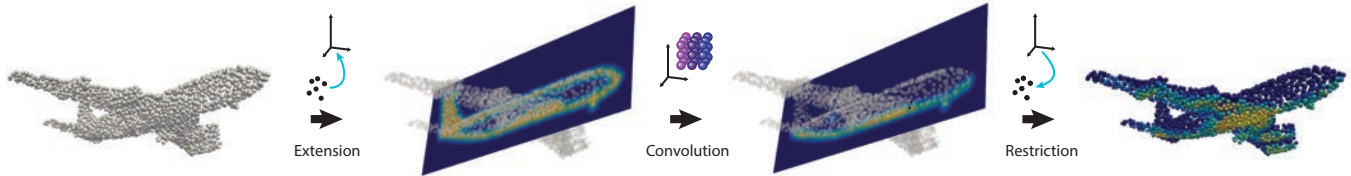


Fig. 1. A new framework for applying convolution to functions defined over point clouds: First, a function over the point cloud (in this case the constant one) is *extended* to a continuous volumetric function over the ambient space; second, a continuous volumetric *convolution* is applied to this function (without any discretization or approximation); and lastly, the result is *restricted* back to the point cloud.

This paper presents Point Convolutional Neural Networks (PCNN): a novel framework for applying convolutional neural networks to point clouds. The framework consists of two operators: extension and restriction, mapping point cloud functions to volumetric functions and vice-versa. A point cloud convolution is defined by pull-back of the Euclidean volumetric convolution via an extension-restriction mechanism.

The point cloud convolution is computationally efficient, invariant to the order of points in the point cloud, robust to different samplings and varying densities, and translation invariant, that is the same convolution kernel is used at all points. PCNN generalizes image CNNs and allows readily adapting their architectures to the point cloud setting.

Evaluation of PCNN on three central point cloud learning benchmarks convincingly outperform competing point cloud learning methods, and the vast majority of methods working with more informative shape representations such as surfaces and/or normals.

CCS Concepts: • **Computing methodologies** → **Neural networks**; **Shape analysis**;

Additional Key Words and Phrases: Shape Analysis, Geometric Deep Learning, Shape Segmentation, Shape Classification, Normal Estimation.

## ACM Reference Format:

Matan Atzmon, Haggai Maron, and Yaron Lipman. 2018. Point Convolutional Neural Networks by Extension Operators. *ACM Trans. Graph.* 37, 4, Article 71 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201301>

## 1 INTRODUCTION

The huge success of deep learning in image analysis motivates researchers to generalize deep learning techniques to work on 3D

shapes. Differently from images, 3D data has several popular representations, most notably surface meshes and points clouds. Surface-based methods exploit connectivity information for 3D deep learning based on rendering [Su et al. 2015], local and global parameterization [Maron et al. 2017; Masci et al. 2015; Sinha et al. 2016], or spectral properties [Yi et al. 2016b]. Point cloud methods rely mostly on points' locations in three-dimensional space and need to implicitly infer how the points are connected to form the underlying shape.

The goal of this paper is to introduce Point Cloud Convolutional Neural Networks (PCNN) generalizing deep learning techniques, and in particular Convolutional Neural Networks (CNN) [Krizhevsky et al. 2012], to point clouds. As a point cloud  $X \subset \mathbb{R}^3$  is merely an approximation to some underlying shape  $S$ , the main challenges in building point cloud networks are to achieve: (i) Invariance to the order of points supplied in  $X$ ; (ii) Robustness to sampling density and distribution of  $X$  in  $S$ ; and (iii) Translation invariance of the convolution operator (*i.e.*, same convolution kernel is used at all points).

Invariance to point order in  $X$  was previously tackled in [Qi et al. 2016a, 2017; Ravanbakhsh et al. 2016; Zaheer et al. 2017] by designing networks that are composition of equivariant layers (*i.e.*, commute with permutations) and a final symmetric layer (*i.e.*, invariant to permutations). As shown in [Ravanbakhsh et al. 2016], any linear equivariant layer is a combination of scaled identity and constant linear operator and therefore missing many of the degrees of freedom existing in standard linear layers such as fully connected and even convolutional.

Volumetric grid methods [Maturana and Scherer 2015; Qi et al. 2016b; Riegler et al. 2016; Wu et al. 2015] use 3D occupancy grid to deal with the point order in  $X$  and provide translation invariance of the convolution operator. However, they quantize the point cloud to a 3D grid, usually producing a crude approximation to the underlying shape (*i.e.*, piecewise constant on voxels) and are confined to a fixed 3D grid structure.

Our approach toward these challenges is to define CNN on a point cloud  $X$  by “pulling back” the standard three dimensional convolution in  $\mathbb{R}^3$  to the point cloud. In more details, given a point

\*Both authors contributed equally to the paper

Authors' addresses: Matan Atzmon, Weizmann Institute of Science; Haggai Maron, Weizmann Institute of Science; Yaron Lipman, Weizmann Institute of Science.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART71 \$15.00

<https://doi.org/10.1145/3197517.3201301>



# Semantic Soft Segmentation

YAĞIZ AKSOY, MIT CSAIL, USA and ETH Zürich, Switzerland

TAE-HYUN OH, MIT CSAIL, USA

SYLVAIN PARIS, Adobe Research, USA

MARC POLLEFEYS, ETH Zürich, Switzerland and Microsoft, USA

WOJCIECH MATUSIK, MIT CSAIL, USA



Fig. 1. We propose a method that can generate soft segments, i.e. layers that represent the semantically meaningful regions as well as the soft transitions between them, automatically by fusing high-level and low-level image features in a single graph structure. The semantic soft segments, visualized by assigning each segment a solid color, can be used as masks for targeted image editing tasks, or selected layers can be used for compositing after layer color estimation. Original images are from [Lin et al. 2014] (top-left, bottom-right), by Death to the Stock Photo (top-right) and by Y. Aksoy (bottom-left).

Accurate representation of soft transitions between image regions is essential for high-quality image editing and compositing. Current techniques for generating such representations depend heavily on interaction by a skilled visual artist, as creating such accurate object selections is a tedious task. In this work, we introduce *semantic soft segments*, a set of layers that correspond to semantically meaningful regions in an image with accurate soft transitions between different objects. We approach this problem from

a spectral segmentation angle and propose a graph structure that embeds texture and color features from the image as well as higher-level semantic information generated by a neural network. The soft segments are generated via eigendecomposition of the carefully constructed Laplacian matrix fully automatically. We demonstrate that otherwise complex image editing tasks can be done with little effort using semantic soft segments.

CCS Concepts: • **Computing methodologies** → **Image segmentation**;

Additional Key Words and Phrases: soft segmentation, semantic segmentation, natural image matting, image editing, spectral segmentation

## ACM Reference Format:

Yağız Aksoy, Tae-Hyun Oh, Sylvain Paris, Marc Pollefeys, and Wojciech Matusik. 2018. Semantic Soft Segmentation. *ACM Trans. Graph.* 37, 4, Article 72 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201275>

## 1 INTRODUCTION

Selection and compositing are at the core of the image editing process. For instance, local adjustments often start with a selection, and combining elements from different images is a powerful way to produce new content. But creating an accurate selection is a tedious task especially when fuzzy boundaries and transparency

Authors' addresses: Yağız Aksoy, MIT CSAIL, 32 Vassar St. Cambridge, MA, 02139, USA, ETH Zürich, Department of Computer Science, Rämistrasse 101, Zürich, 8092, Switzerland; Tae-Hyun Oh, MIT CSAIL, 32 Vassar St. Cambridge, MA, 02139, USA; Sylvain Paris, Adobe Research, One Broadway, Cambridge, MA, 02142, USA; Marc Pollefeys, ETH Zürich, Department of Computer Science, Rämistrasse 101, Zürich, 8092, Switzerland, Microsoft, Redmond, WA, USA; Wojciech Matusik, MIT CSAIL, 32 Vassar St. Cambridge, MA, 02139, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART72 \$15.00 <https://doi.org/10.1145/3197517.3201275>

# Efficient Rendering of Layered Materials using an Atomic Decomposition with Statistical Operators

LAURENT BELCOUR, Unity Technologies



Fig. 1. We introduce a BSDF model to render plane parallel layered materials using an analysis of the directional statistics of light interaction with microfacet geometry and participating media. Our model closely matches the reference and supports an arbitrary number of textured layers while being energy conserving, free from heavy per-material precomputation, and compatible with real-time constraints.

We derive a novel framework for the efficient analysis and computation of light transport within layered materials. Our derivation consists in two steps. First, we decompose light transport into a set of atomic operators that act on its directional statistics. Specifically, our operators consist of reflection, refraction, scattering, and absorption, whose combinations are sufficient to describe the statistics of light scattering multiple times within layered structures. We show that the first three directional moments (energy, mean and variance) already provide an accurate summary. Second, we extend the adding-doubling method to support arbitrary combinations of such operators efficiently. During shading, we map the directional moments to BSDF lobes. We validate that the resulting BSDF closely matches the ground truth in a lightweight and efficient form. Unlike previous methods we support an arbitrary number of textured layers, and demonstrate a practical and accurate rendering of layered materials with both an offline and real-time implementation that are free from per-material precomputation.

CCS Concepts: • **Computing methodologies** → **Reflectance modeling**;

Additional Key Words and Phrases: Layered Materials, Statistical Analysis, Real-Time Rendering

Author's address: Laurent Belcour, Unity Technologies, 41 avenue de l'Europe, 38330 Montbonnot Saint-Martin, FRANCE, laurent@unity3d.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART73 \$15.00 <https://doi.org/10.1145/3197517.3201289>

## ACM Reference Format:

Laurent Belcour. 2018. Efficient Rendering of Layered Materials using an Atomic Decomposition with Statistical Operators. *ACM Trans. Graph.* 37, 4, Article 73 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201289>

## 1 INTRODUCTION

Elaborate shading models allow the realistic reproduction of the diversity and complexity of surface appearance found in nature or man-made objects. A simple way to enrich a set of materials is to coat them with one or many strata of glazing or clear coat. This is typically what happens when objects are wet or paints are glazed. For example applying a clear coat on a rough metal will create a distinctive *hazy* look. Unfortunately, coatings increase the difficulty of modeling the resulting appearance and design efficient shading models for rendering. Up to now, nobody provided a comprehensive model of the *Bidirectional Scattering Distribution Function* (BSDF) of coated materials. In this work, we focus on a subset of coatings called *layered materials* [Jakob et al. 2014].

**Layered Materials.** Layered materials assume that the coating layers are plane-parallel with respect to the shading normal separated by rough interfaces. It is also assumed that geometrical optics fully describes the appearance. Thus, the thicknesses of layers is larger than the wavelengths of visible light (removing wave optics from our study). However, the number of layers can be arbitrary and thus light can bounce many times before exiting the material, which makes evaluation difficult. Offline productions [Hery et al. 2017; Kulla and Conty 2017; Langlands 2014] and the real-time community [Drobot 2017] have expressed interest in those appearance. However, they use a heavy restriction of layered materials that consist of only two layers evaluated with single reflection per layer (that

# The Layer Laboratory: A Calculus for Additive and Subtractive Composition of Anisotropic Surface Reflectance

TIZIAN ZELTNER and WENZEL JAKOB, École polytechnique fédérale de Lausanne (EPFL)



Fig. 1. All materials in this rendering of an interior scene were generated using the techniques proposed in this article. Insets on the left side reveal the corresponding layer structures, which make use of dielectric and conductive microfacet models, scattering and absorbing volumes, and measured materials acquired using dense goniophotometric measurements.

We present a versatile computational framework for modeling the reflective and transmissive properties of arbitrarily layered anisotropic material structures. Given a set of input layers, our model synthesizes an effective BSDF of the entire structure, which accounts for all orders of internal scattering and is efficient to sample and evaluate in modern rendering systems.

Our technique builds on the insight that reflectance data is sparse when expanded into a suitable frequency-space representation, and that this property extends to the class of anisotropic materials. This sparsity enables an efficient matrix calculus that admits the entire space of BSDFs and considerably expands the scope of prior work on layered material modeling. We show how both measured data and the popular class of microfacet models

can be expressed in our representation, and how the presence of anisotropy leads to a weak coupling between Fourier orders in frequency space.

In addition to additive composition, our models support subtractive composition, a fascinating new operation that reconstructs the BSDF of a material that can only be observed indirectly through another layer with known reflectance properties. The operation produces a new BSDF of the desired layer as if measured in isolation. Subtractive composition can be interpreted as a type of deconvolution that removes both internal scattering and blurring due to transmission through the known layer.

We experimentally demonstrate the accuracy and scope of our model and validate both additive and subtractive composition using measurements of real-world layered materials. Both implementation and data will be released to ensure full reproducibility of all of our results.<sup>1</sup>

Authors' address: Tizian Zeltner; Wenzel Jakob, École polytechnique fédérale de Lausanne (EPFL).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART74 \$15.00  
<https://doi.org/10.1145/3197517.3201321>

CCS Concepts: • **Computing methodologies** → **Reflectance modeling**;

Additional Key Words and Phrases: Layered materials, adding equations, subtracting equations, BRDF, BSDF

## ACM Reference Format:

Tizian Zeltner and Wenzel Jakob. 2018. The Layer Laboratory: A Calculus for Additive and Subtractive Composition of Anisotropic Surface Reflectance. *ACM Trans. Graph.* 37, 4, Article 74 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201321>

<sup>1</sup> Available at <http://rgl.epfl.ch/publications/Zeltner2018Layer>



# Rendering Specular Microgeometry with Wave Optics

LING-QI YAN, University of California, Berkeley

MILOŠ HAŠAN, Autodesk

BRUCE WALTER, Cornell University

STEVE MARSCHNER, Cornell University

RAVI RAMAMOORTHY, University of California, San Diego

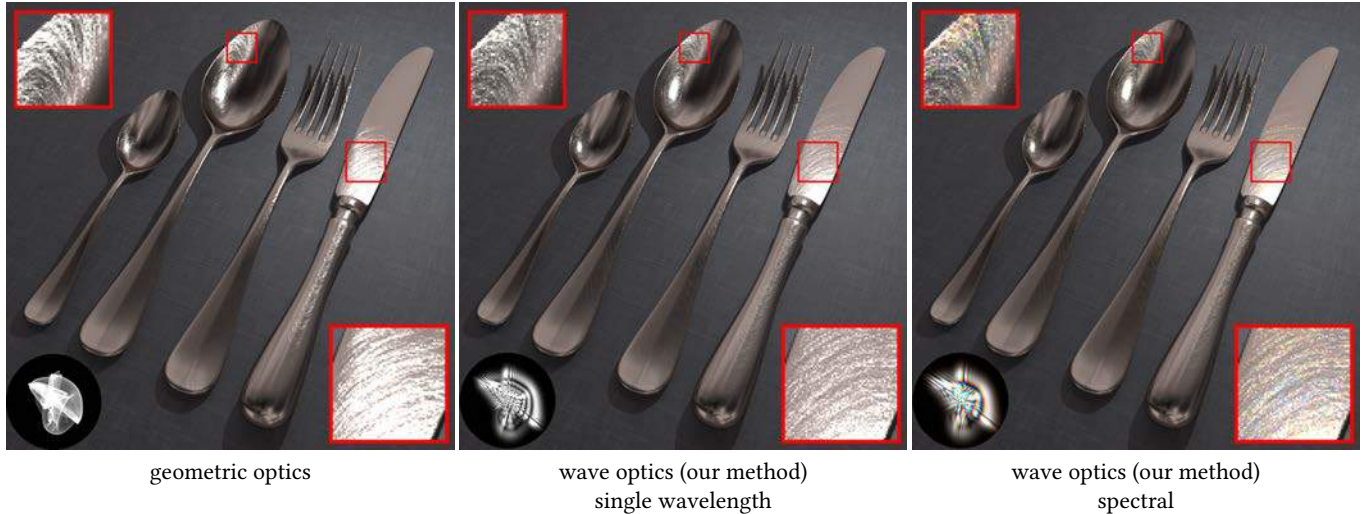


Fig. 1. We present the first practical method for rendering specular reflection from arbitrary high-resolution microstructure (represented as discretized heightfields) using wave optics. Left: Rendering with previous work [Yan et al. 2016], based on the rules of geometric optics. Middle: Using wave optics, even with a single fixed wavelength, our method generates a more natural appearance as compared to geometric optics. Right: A spectral rendering additionally shows subtle but important color glint effects. Insets show enlarged regions and representative BRDFs generated using each method. We encourage readers to zoom in to better see color and detail, and to view the full resolution supplementary images to see the subtle details in all of the figures.

Simulation of light reflection from specular surfaces is a core problem of computer graphics. Existing solutions either make the approximation of providing only a large-area average solution in terms of a fixed BRDF (ignoring spatial detail), or are specialized for specific microgeometry (e.g. 1D scratches), or are based only on geometric optics (which is an approximation to more accurate wave optics). We design the first rendering algorithm based on a wave optics model that is also able to compute spatially-varying specular highlights with high-resolution detail on general surface microgeometry. We compute a wave optics reflection integral over the coherence area; our solution is based on approximating the phase-delay grating representation

Authors' addresses: Ling-Qi Yan, University of California, Berkeley, lingqi@berkeley.edu; Miloš Hašan, Autodesk, milos.hasan@gmail.com; Bruce Walter, Cornell University, bruce.walter@cornell.edu; Steve Marschner, Cornell University, srm@cs.cornell.edu; Ravi Ramamoorthi, University of California, San Diego, ravir@cs.ucsd.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART75 \$15.00  
<https://doi.org/10.1145/3197517.3201351>

of a micron-resolution surface heightfield using Gabor kernels. We found that the appearance difference between the geometric and wave solution is more dramatic when spatial detail is taken into account. The visualizations of the corresponding BRDF lobes differ significantly. Moreover, the wave optics solution varies as a function of wavelength, predicting noticeable color effects in the highlights. Our results show both single-wavelength and spectral solution to reflection from common everyday objects, such as brushed, scratched and bumpy metals.

CCS Concepts: • **Computing methodologies** → **Rendering**;

Additional Key Words and Phrases: specular surface rendering, glints, material appearance, Harvey-Shack, wave optics

## ACM Reference Format:

Ling-Qi Yan, Miloš Hašan, Bruce Walter, Steve Marschner, and Ravi Ramamoorthi. 2018. Rendering Specular Microgeometry with Wave Optics. *ACM Trans. Graph.* 37, 4, Article 75 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201351>

## 1 INTRODUCTION

Simulation of material appearance is a core problem of computer graphics, and specular highlight appearance is among the most common effects. One of the most fundamental questions is: given a

# Gaussian Material Synthesis

KÁROLY ZSOLNAI-FEHÉR, TU Wien

PETER WONKA, KAUST

MICHAEL WIMMER, TU Wien

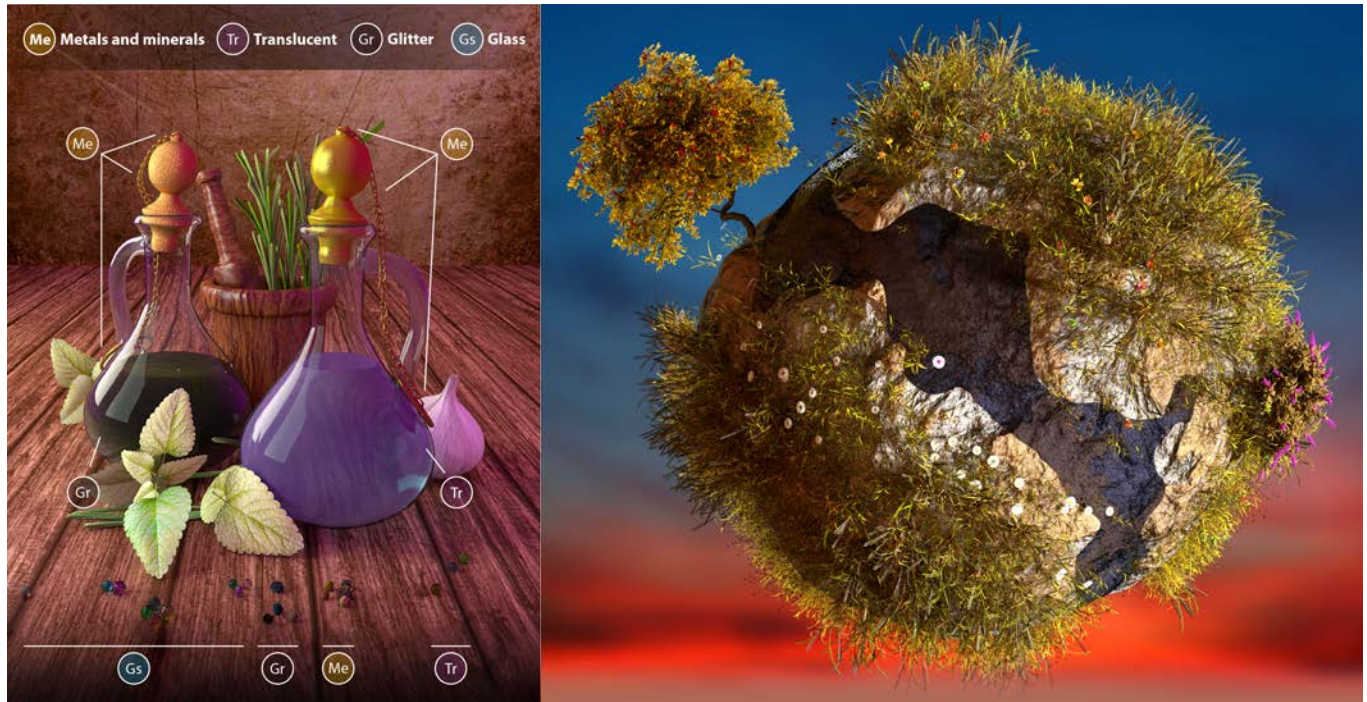


Fig. 1. Our system opens up the possibility of rapid mass-scale material synthesis for novice and expert users alike. This method takes a set of user preferences as an input and recommends relevant new materials from the learned distributions. On the left, we populated a scene with metals and minerals, translucent, glittery and glassy materials, each of which was learned and synthesized via our proposed technique. The image on the right showcases rich material variations for more than a hundred synthesized materials and objects for the vegetation of the planet. The learning and recommendation steps take less than a minute.

We present a learning-based system for rapid mass-scale material synthesis that is useful for novice and expert users alike. The user preferences are learned via Gaussian Process Regression and can be easily sampled for new recommendations. Typically, each recommendation takes 40-60 seconds to render with global illumination, which makes this process impracticable for real-world workflows. Our neural network eliminates this bottleneck by providing high-quality image predictions in real time, after which it is possible to pick the desired materials from a gallery and assign them to a scene in an intuitive manner. Workflow timings against Disney's "principled" shader reveal that our system scales well with the number of sought materials,

Authors' addresses: Károly Zsolnai-Fehér, TU Wien, Favoritenstrasse 9-11/193-02, Vienna, Austria, 1040, zsolnai@cg.tuwien.ac.at; Peter Wonka, KAUST, Al Khwarizmi Bldg 1, Thuwal, Kingdom of Saudi Arabia, 23955-6900, pwonka@gmail.com; Michael Wimmer, TU Wien, Favoritenstrasse 9-11/193-02, Vienna, Austria, 1040, wimmer@cg.tuwien.ac.at.

ACM acknowledges that this contribution was authored or co-authored by an employee, contractor, or affiliate of the United States government. As such, the United States government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for government purposes only.

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART6

<https://doi.org/10.1145/3197517.3201307>

thus empowering even novice users to generate hundreds of high-quality material models without any expertise in material modeling. Similarly, expert users experience a significant decrease in the total modeling time when populating a scene with materials. Furthermore, our proposed solution also offers controllable recommendations and a novel latent space variant generation step to enable the real-time fine-tuning of materials without requiring any domain expertise.

CCS Concepts: • **Computing methodologies** → **Gaussian processes; Neural networks; Rendering; Ray tracing;**

Additional Key Words and Phrases: neural networks, photorealistic rendering, gaussian process regression, neural rendering, latent variables

## ACM Reference Format:

Károly Zsolnai-Fehér, Peter Wonka, and Michael Wimmer. 2018. Gaussian Material Synthesis. *ACM Trans. Graph.* 37, 4, Article 76 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201307>

## 1 INTRODUCTION

Light transport simulations are the industry standard way to create high-quality photorealistic imagery. This class of techniques



# Developability of Triangle Meshes

ODED STEIN and EITAN GRINSPUN, Columbia University  
KEENAN CRANE, Carnegie Mellon University

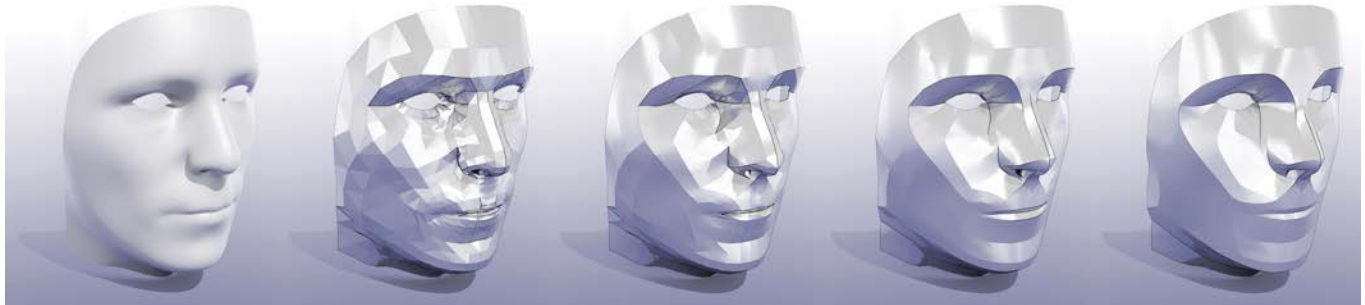


Fig. 1. By encouraging discrete developability, a given mesh evolves toward a shape comprised of flattenable pieces separated by highly regular seam curves.

*Developable surfaces* are those that can be made by smoothly bending flat pieces without stretching or shearing. We introduce a definition of developability for triangle meshes which exactly captures two key properties of smooth developable surfaces, namely flattenability and presence of straight ruling lines. This definition provides a starting point for algorithms in developable surface modeling—we consider a variational approach that drives a given mesh toward developable pieces separated by regular seam curves. Computation amounts to gradient descent on an energy with support in the vertex star, without the need to explicitly cluster patches or identify seams. We briefly explore applications to developable design and manufacturing.

CCS Concepts: • **Mathematics of computing** → **Discretization**; • **Computing methodologies** → **Mesh geometry models**;

Additional Key Words and Phrases: developable surface modeling, discrete differential geometry, digital geometry processing

## ACM Reference Format:

Oded Stein, Eitan Grinspun, and Keenan Crane. 2018. Developability of Triangle Meshes. *ACM Trans. Graph.* 37, 4, Article 77 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201303>

## 1 INTRODUCTION

Fabrication from developable pieces provides an enticing paradigm for manufacturing: flat sheet materials such as plywood or sheet metal are easy to cut, ship, and store; surfaces comprised of developable pieces also reduce cost and improve quality in computer controlled milling [Harik et al. 2013]. To date, however, there are few tools for automatic conversion of curved surfaces into developable

pieces—most industrial applications still rely on manual interaction and designer expertise [Chang 2015].

The goal of this paper is to develop mathematical and computational foundations for developability in the simplicial setting, and show how this perspective inspires new approaches to developable design. Our starting point is a new definition of developability for triangle meshes (Section 3). This definition is motivated by the behavior of developable surfaces that are twice differentiable rather than those that are merely continuous: instead of just asking for flattenability, we also seek a definition that naturally leads to well-defined *ruling lines*. Moreover, unlike existing notions of discrete developability, it applies to general triangulated surfaces, with no special conditions on combinatorics.

These features make our definition a natural starting point for algorithms that seek to design developable surfaces. In this paper, we investigate a global variational approach that encourages the discrete developability of each vertex (Section 4.1). An interesting observation is that this local optimization naturally tends toward surfaces that are *piecewise* developable: in practice, curvature concentrates onto a sparse collection of seam curves (Section 4.1.3) which are themselves highly regular (Section 4.2).

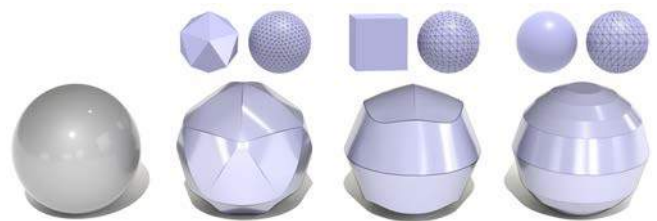


Fig. 2. How should one approximate a perfect sphere by developable pieces? Due to symmetry, there is no canonical “best” answer. Instead, we guide the surface toward a desired design by choosing an initial tessellation (shown at top) that breaks this symmetry.

Authors’ addresses: Oded Stein, Eitan Grinspun, Columbia University, 530 West 120th Street, New York, NY, 10027; Keenan Crane, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART77 \$15.00

<https://doi.org/10.1145/3197517.3201303>



# Shape Representation by Zippables

CHRISTIAN SCHÜLLER, ROI PORANNE, and OLGA SORKINE-HORNUNG, ETH Zurich, Switzerland

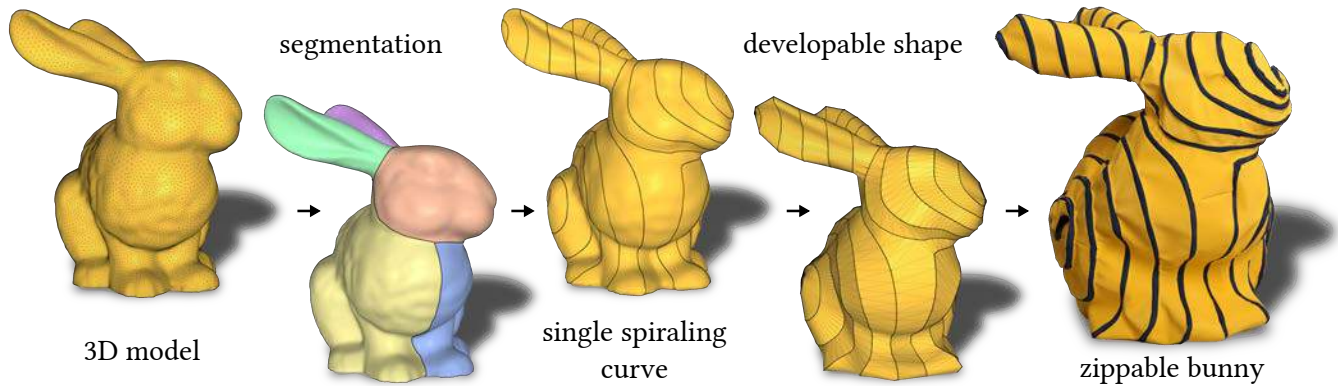


Fig. 1. The pipeline of our approach. Starting from a 3D model, the user decomposes the shape into topological cylinders. Our algorithm automatically produces a single continuous curve on the shape that spirals along the cylinders. It proceeds to cut the shape along the curve and creates a developable surface that can be trivially unfolded into a single 2D shape – the so called *zippable*. Based on the flattening, plans for laser cutting it from fabric are generated. Finally, we attach a zipper with a single slider to the boundary of the zippable. Zipping it up reproduces a faithful approximation of the input model.

Fabrication from developable parts is the basis for arts such as papercraft and needlework, as well as modern architecture and CAD in general, and it has inspired much research. We observe that the assembly of complex 3D shapes created by existing methods often requires first fabricating many small parts and then carefully following instructions to assemble them together. Despite its significance, this error prone and tedious process is generally neglected in the discussion. We present the concept of *zippables* – single, two dimensional, branching, ribbon-like pieces of fabric that can be quickly zipped up without any instructions to form 3D objects. Our inspiration comes from the so-called *zipit* bags [zipit 2017], which are made of a single, long ribbon with a zipper around its boundary. In order to “assemble” the bag, one simply needs to zip up the ribbon. Our method operates in the same fashion, but it can be used to approximate a wide variety of shapes. Given a 3D model, our algorithm produces plans for a single 2D shape that can be laser cut in few parts from fabric or paper. A zipper can then be attached along the boundary by sewing, or by gluing using a custom-built fastening rig. We show physical and virtual results that demonstrate the capabilities of our method and the ease with which shapes can be assembled.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Shape modeling**;

Additional Key Words and Phrases: Digital fabrication, zippers, developable surfaces, mesh parameterization

This work was supported in part by the ERC grant iModel (StG-2012-306877). Authors’ address: Christian Schüller; Roi Poranne; Olga Sorkine-Hornung, ETH Zurich, Department of Computer Science, Universitätsstrasse 6, Zurich, 8092, Switzerland.



This work is licensed under a Creative Commons Attribution-ShareAlike International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART78  
<https://doi.org/10.1145/3197517.3201347>

## ACM Reference Format:

Christian Schüller, Roi Poranne, and Olga Sorkine-Hornung. 2018. Shape Representation by Zippables. *ACM Trans. Graph.* 37, 4, Article 78 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201347>

## 1 INTRODUCTION

Representing shapes using developable surfaces is a problem with numerous applications, ranging from recreational activities like papercraft and plush toy fabrication, to large scale industrial design and modern architecture. In this paper we introduce the concept of



Fig. 2. The star model fabricated with our fastening rig. The insets on the left and in the middle show the developable model from a front and side perspective. Note how the fabricated star perfectly resembles the predicted shape. The segmentation is shown in the top right corner; below it a visualization of the zipper tape and the flattened *zippable*. The physical result has a height of 27 cm.

# FEPR: Fast Energy Projection for Real-Time Simulation of Deformable Objects

DIMITAR DINEV\*, University of Utah

TIANTIAN LIU\*, University of Pennsylvania

JING LI, University of Utah

BERNHARD THOMASZEWSKI, Université de Montréal

LADISLAV KAVAN, University of Utah

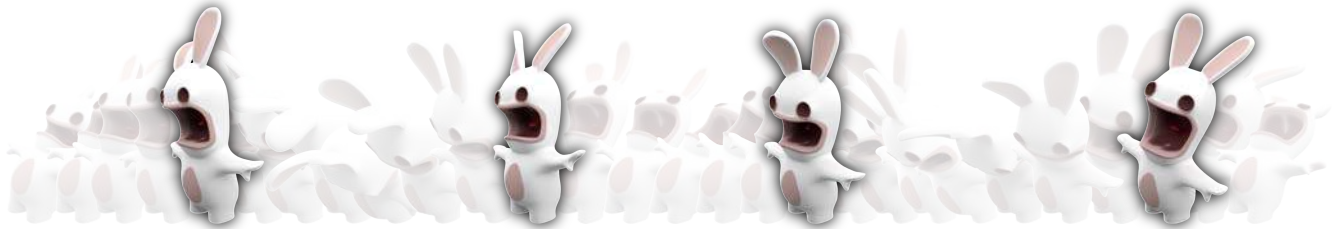


Fig. 1. Our fast energy projection method produces vivid motion for an animated rabbit.

We propose a novel projection scheme that corrects energy fluctuations in simulations of deformable objects, thereby removing unwanted numerical dissipation and numerical “explosions”. The key idea of our method is to first take a step using a conventional integrator, then project the result back to the constant energy-momentum manifold. We implement this strategy using *fast projection*, which only adds a small amount of overhead to existing physics-based solvers. We test our method with several implicit integration rules and demonstrate its benefits when used in conjunction with Position Based Dynamics and Projective Dynamics. When added to a dissipative integrator such as backward Euler, our method corrects the artificial damping and thus produces more vivid motion. Our projection scheme also effectively prevents instabilities that can arise due to approximate solves or large time steps. Our method is fast, stable, and easy to implement—traits that make it well-suited for real-time physics applications such as games or training simulators.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: Physics-based animation, real-time simulation.

## ACM Reference Format:

Dimitar Dinev, Tiantian Liu, Jing Li, Bernhard Thomaszewski, and Ladislav Kavan. 2018. FEPR: Fast Energy Projection for Real-Time Simulation of

\*The first two authors contributed equally to this work.

Authors’ addresses: Dimitar Dinev, University of Utah; Tiantian Liu, University of Pennsylvania; Jing Li, University of Utah; Bernhard Thomaszewski, Université de Montréal; Ladislav Kavan, University of Utah.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART79 \$15.00

<https://doi.org/10.1145/3197517.3201277>

Deformable Objects. *ACM Trans. Graph.* 37, 4, Article 79 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201277>

## 1 INTRODUCTION

Simulation of dynamic elastic objects is important in many areas of engineering where one of the key numerical considerations is accuracy. Consequently, many numerical integration techniques have been developed with an often implicitly assumed understanding that accuracy is paramount. This is perfectly justified in engineering applications when the goal is to simulate the dynamics of structures such as a bridge or an airplane wing. Quite different situations occur in interactive physics-based simulations, important e.g. in computer games or surgical simulators [Sui et al. 2017]. The key difference is that numerical time integration must be fast enough to provide solutions in real-time, which often means frame rates of 30Hz or more. The simulation results are immediately displayed to the user and therefore, the numerical methods must be stable, because it is unacceptable to display artifacts such as erroneously “exploded” frames to the user. With fast advances in technologies for virtual and augmented reality, we can expect increasing demands for real-time physics simulators.

Due to this intolerance to instabilities, real-time simulators typically rely on dissipative integrators such as backward Euler. However, these integrators come with two types of drawbacks: 1) the increased stability comes at the cost of artificial numerical damping which depends on the time step size and other parameters of the simulated system; 2) implicit integration rules are not easy to solve numerically. In theory, machine-precision-accurate solutions of implicit update rules can be computed, e.g., by iterating Newton’s method to convergence. However, in practice, the implicit integration rules are often only solved approximatively, e.g., by linearizing the nonlinear implicit equations [Baraff and Witkin 1998]. There are also methods developed specifically for real-time physics, such

# Hyper-Reduced Projective Dynamics

CHRISTOPHER BRANDT, ELMAR EISEMANN, and KLAUS HILDEBRANDT, Delft University of Technology, The Netherlands

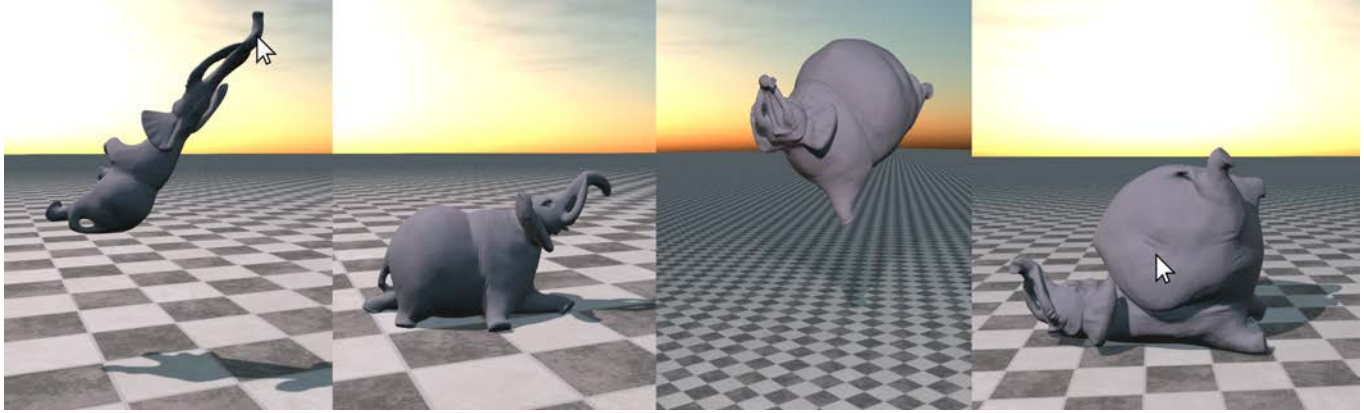


Fig. 1. Our Hyper-Reduced Projective Dynamics method is able to handle different types of constraints simultaneously (volume preservation for tetrahedrons and strain resistance for boundary triangles) for complex geometries (52k vertices) in real-time. During the simulation we change the target volume of the tetrahedrons, which causes a balloon-like blowup effect. We resolve ground collisions and allow user interaction by click and drag. The simulation runs at 62 fps including rendering. The total precomputation time of our method is 21 seconds.

We present a method for the real-time simulation of deformable objects that combines the robustness, generality, and high performance of Projective Dynamics with the efficiency and scalability offered by model reduction techniques. The method decouples the cost for time integration from the mesh resolution and can simulate large meshes in real-time. The proposed hyper-reduction of Projective Dynamics combines a novel fast approximation method for constraint projections and a scalable construction of sparse subspace bases. The resulting system achieves real-time rates for large subspaces enabling rich dynamics and can resolve general user interactions, collision constraints, external forces and changes to the materials. The construction of the hyper-reduced system does not require user-interaction and refrains from using training data or modal analysis, which results in a fast preprocessing stage.

CCS Concepts: • **Computing methodologies** → **Real-time simulation**; **Physical simulation**;

Additional Key Words and Phrases: Real-time simulation, projective dynamics, model reduction, subspace dynamics, reduced-order model

Authors' address: Christopher Brandt, c.brandt@tudelft.nl; Elmar Eiseemann, e.eiseemann@tudelft.nl; Klaus Hildebrandt, k.a.hildebrandt@tudelft.nl, Delft University of Technology, Department of Intelligent Systems, Van Mourik Broekmanweg 6, Delft, 2628 XE, The Netherlands.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART80 \$15.00

<https://doi.org/10.1145/3197517.3201387>

## ACM Reference Format:

Christopher Brandt, Elmar Eiseemann, and Klaus Hildebrandt. 2018. Hyper-Reduced Projective Dynamics. *ACM Trans. Graph.* 37, 4, Article 80 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201387>

## 1 INTRODUCTION

The simulation of deformable objects is an important factor in various entertainment and training applications of computer graphics. Of particular interest are real-time simulations since these allow the user to interact with the simulation and thereby greatly enrich the experience in games, virtual reality, artistic applications and medical training. The combination of the nonlinear nature of deformable objects and the geometric complexity of objects in these scenarios make real-time simulation a challenging problem. In addition to efficiency, robustness is important for real-time simulation since interactivity leads to unforeseeable states of the system.

Projective Dynamics, introduced in [Bouaziz et al. 2014; Liu et al. 2013], is a simulation framework that is *general*, since it allows for the simulation of different types of deformable objects (solids, shells and rods) and materials. It is *robust*, as it is capable of adequately handling large time steps, large deformations and degenerate geometries, and it is *fast*. On the technical level, a variational implicit time integration scheme is used and the resulting optimization problem is solved by an alternating local/global approach. The local steps in the optimization can be executed in parallel, and for *constraint-based* potentials, the system matrix of the linear system to be solved in the global steps is constant, allowing for the re-use of a sparse factorization of the matrix. While this makes Projective Dynamics highly efficient for real-time simulation, it is still limited to objects that can be represented by coarse geometric models.



# Dynamic Kelvinlets: Secondary Motions based on Fundamental Solutions of Elastodynamics

FERNANDO DE GOES, Pixar Animation Studios

DOUG L. JAMES, Pixar Animation Studios and Stanford University

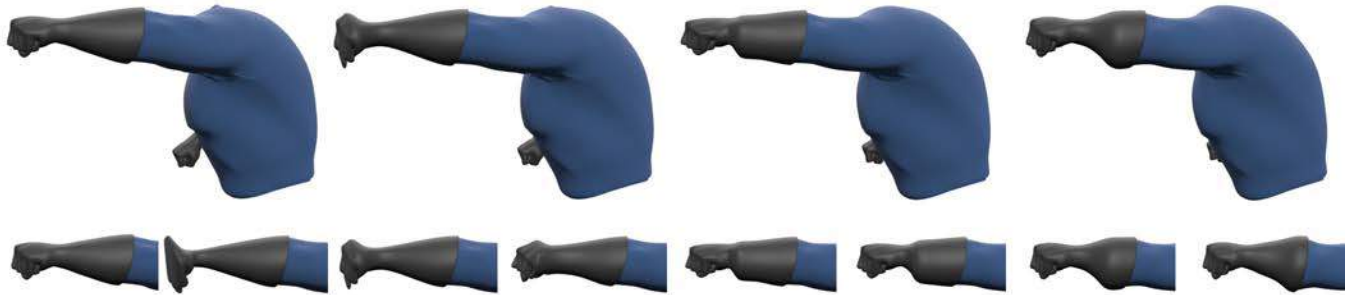


Fig. 1. **Dynamic Kelvinlets:** A punch animation is augmented with impact waves generated by dynamic Kelvinlets, a novel approach for computing volumetric deformations with real-time feedback based on analytical solutions of linear elastodynamics. Top row shows elastic waves propagated across the character's fist and forearm after impact with an imaginary wall. Bottom row shows a close-up of the animation sequence. ©Disney/Pixar

We introduce *Dynamic Kelvinlets*, a new analytical technique for real-time physically based animation of virtual elastic materials. Our formulation is based on the dynamic response to time-varying force distributions applied to an infinite elastic medium. The resulting displacements provide the plausibility of volumetric elasticity, the dynamics of compressive and shear waves, and the interactivity of closed-form expressions. Our approach builds upon the work of de Goes and James [2017] by presenting an extension of the regularized Kelvinlet solutions from elastostatics to the elastodynamic regime. To finely control our elastic deformations, we also describe the construction of compound solutions that resolve pointwise and keyframe constraints. We demonstrate the versatility and efficiency of our method with a series of examples in a production grade implementation.

CCS Concepts: • **Computing methodologies** → **Physical simulation; Procedural animation;**

Additional Key Words and Phrases: elastic waves, linear elastodynamics.

## ACM Reference Format:

Fernando de Goes and Doug L. James. 2018. Dynamic Kelvinlets: Secondary Motions based on Fundamental Solutions of Elastodynamics. *ACM Trans. Graph.* 37, 4, Article 81 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201280>

## 1 INTRODUCTION

Physically based simulation is a mainstay of computer graphics widely used to animate virtual characters and natural phenomena. Despite their success, existing physics-driven techniques still have

several downsides that interrupt the animation process. For instance, the dynamics computations are cumbersome due to numerical solves and time-stepping conditions. Another major impediment is the tedious setup phase which often involves volumetric meshing. These limitations motivate digital artists to use simpler models that capture the visual cues at interactive rates, especially for small-scale and secondary deformations where approximated solutions may suffice.

In this paper, we introduce a new interactive tool for the generation of physically based dynamics with elastic deformations. Our approach is based on the extension of the elastostatic regularized Kelvinlets [de Goes and James 2017] to dynamics. This is achieved by deriving novel fundamental solutions of elastodynamics for spatially regularized and time-varying forces applied to an infinite continuum. The resulting *dynamic Kelvinlets* lead to analytical closed-form expressions that define wave-like volumetric displacements. Consequently, we can animate deformable models free of any geometric discretization, computationally intensive solve, or stability restriction. Instead, our elastic deformations can be evaluated rapidly both in space and time, with interactive control of wave speed and volume compression. While not intended for general simulation tasks, dynamic Kelvinlets are well-suited to procedural visual effects such as jiggling, ripples, denting, and blasts (Figure 1).

Our contributions encompass several types of dynamic Kelvinlets. The first and most fundamental solution is the *impulse dynamic Kelvinlet*, which describes volumetric waves caused by an impulse load. Based on this solution, we derive the elastic response to steady push-like forces, which we call the *push dynamic Kelvinlet*. The latter produces permanent deformations that converge in the quasi-static regime to the regularized Kelvinlets [de Goes and James 2017]. We also present time-varying deformations generated by affine loads such as twist, scale, and pinch, and the construction of constrained and compounded elastic waves.

Authors' addresses: Fernando de Goes, Pixar Animation Studios, [fernando@pixar.com](mailto:fernando@pixar.com); Doug L. James, Pixar Animation Studios, Stanford University, [djames@pixar.com](mailto:djames@pixar.com).

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART81

<https://doi.org/10.1145/3197517.3201280>

# Gradient-domain Volumetric Photon Density Estimation

ADRIEN GRUSON, The University of Tokyo and JFLI CNRS UMI 3527

BINH-SON HUA, The University of Tokyo and Singapore University of Technology and Design

NICOLAS VIBERT, McGill University

DEREK NOWROUZEZHAI, McGill University

TOSHIYA HACHISUKA, The University of Tokyo

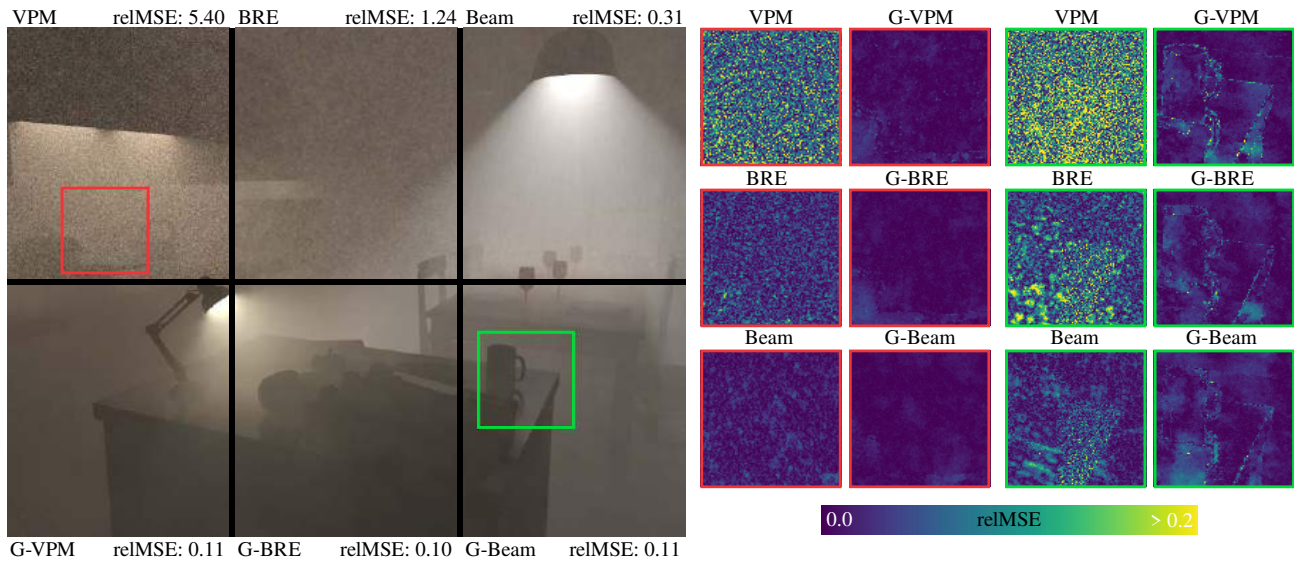


Fig. 1. Equal-time (5 minutes) renderings of a smoky KITCHEN scene. Gradient-domain volumetric rendering techniques with L1 reconstruction converge faster than primal-domain volumetric rendering technique. The relMSE error metric has a unitless scale of  $10^{-2}$ .

Gradient-domain rendering can improve the convergence of surface-based light transport by exploiting smoothness in image space. Scenes with participating media exhibit similar smoothness and could potentially benefit from gradient-domain techniques. We introduce the first gradient-domain formulation of image synthesis with homogeneous participating media, including four novel and efficient gradient-domain volumetric density estimation algorithms. We show that naïve extensions of gradient domain path-space and density estimation methods to volumetric media, while functional, can result in inefficient estimators. Focussing on point-, beam- and plane-based gradient-domain estimators, we introduce a novel shift mapping that eliminates redundancies in the naïve formulations using spatial relaxation within the volume. We show that gradient-domain volumetric rendering improve convergence compared to primal domain state-of-the-art, across a suite of scenes. Our formulation and algorithms support progressive estimation and are easy to incorporate atop existing renderers.

Authors' addresses: Adrien Gruson, [adrien.gruson@gmail.com](mailto:adrien.gruson@gmail.com); Binh-Son Hua, [binhson.hua@gmail.com](mailto:binhson.hua@gmail.com); Nicolas Vibert, [nicolas.vibert@mail.mcgill.ca](mailto:nicolas.vibert@mail.mcgill.ca); Derek Nowrouzezahrai, [derek@cim.mcgill.ca](mailto:derek@cim.mcgill.ca); Toshiya Hachisuka, [thachisuka@siggraph.org](mailto:thachisuka@siggraph.org).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART77 \$15.00

<https://doi.org/10.1145/3197517.3201363>

CCS Concepts: • **Computing methodologies** → **Ray tracing**;

Additional Key Words and Phrases: gradient rendering, participating media

## ACM Reference Format:

Adrien Gruson, Binh-Son Hua, Nicolas Vibert, Derek Nowrouzezahrai, and Toshiya Hachisuka. 2018. Gradient-domain Volumetric Photon Density Estimation. *ACM Trans. Graph.* 37, 4, Article 77 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201363>

## 1 INTRODUCTION

Accurately simulating light transport in participating media remains a longstanding challenge in computer graphics. While Monte Carlo methods robustly render a range of surface-only scenes, density estimation techniques tend towards more desirable convergence behavior in scenes with volumetric media. Recent advances in volumetric density estimation, such as the beam radiance estimate [Jarosz et al. 2008], photon point and beam estimators [Jarosz et al. 2011a], virtual ray and beam lights [Novák et al. 2012a,b], and other higher-dimensional representation such as photon planes and volumes [Bitterli and Jarosz 2017], further improve this behavior.

Several recent extensions of conventional rendering methods to the gradient domain have greatly improved visual convergence [Ketunen et al. 2015; Lehtinen et al. 2013; Manzi et al. 2015]. The seminal work on gradient-domain Metropolis light transport [Lehtinen et al. 2013] led to gradient-domain variants of uni- and bi-directional path

# A Radiative Transfer Framework for Spatially-Correlated Materials

ADRIAN JARABO, Universidad de Zaragoza – I3A

CARLOS ALIAGA, Universidad de Zaragoza – I3A, and Desilico Labs

DIEGO GUTIERREZ, Universidad de Zaragoza – I3A

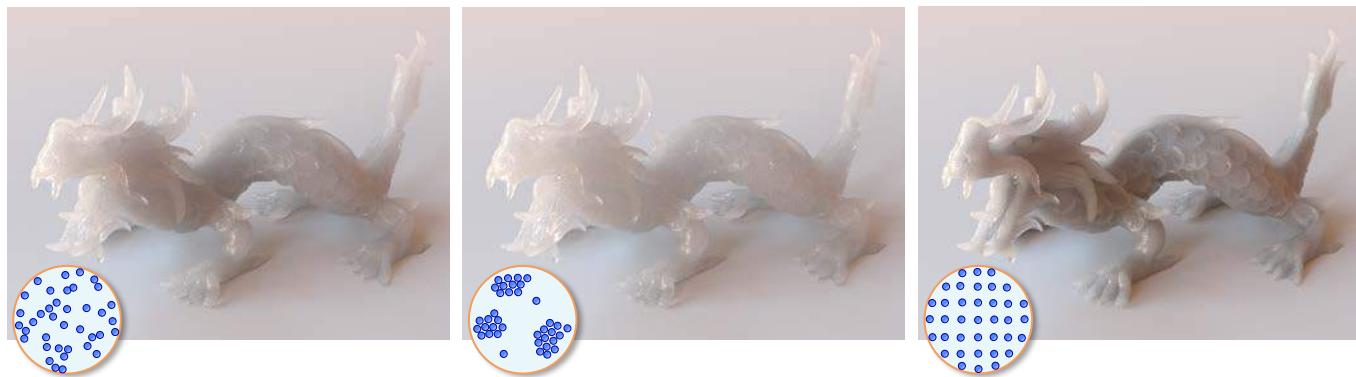


Fig. 1. Spatial correlation in media leads to non-exponential light transport, which significantly affects appearance. The image shows volumetric renderings of translucent dragons made of materials with the same density per unit differential volume  $\bar{\mu} = 10$  (isotropic, albedo  $\Lambda = .8$ ), but different correlations. Left: using classic light transport, where material particles are assumed to be uncorrelated. Middle and right: positive and negative correlation, respectively, using our novel framework for spatially-correlated materials. The insets show illustrative views of scatterer correlation for each dragon.

We introduce a non-exponential radiative framework that takes into account the local spatial correlation of scattering particles in a medium. Most previous works in graphics have ignored this, assuming uncorrelated media with a uniform, random local distribution of particles. However, positive and negative correlation lead to slower- and faster-than-exponential attenuation respectively, which cannot be predicted by the Beer-Lambert law. As our results show, this has a major effect on extinction, and thus appearance. From recent advances in neutron transport, we first introduce our Extended Generalized Boltzmann Equation, and develop a general framework for light transport in correlated media. We lift the limitations of the original formulation, including an analysis of the boundary conditions, and present a model suitable for computer graphics, based on optical properties of the media and statistical distributions of scatterers. In addition, we present an analytic expression for transmittance in the case of positive correlation, and show how to incorporate it efficiently into a Monte Carlo renderer. We show results with a wide range of both positive and negative correlation, and demonstrate the differences compared to classic light transport.

CCS Concepts: • **Computing methodologies** → **Rendering**; *Reflectance modeling*;

Authors' addresses: Adrian Jarabo, Universidad de Zaragoza – I3A, [ajarabo@unizar.es](mailto:ajarabo@unizar.es); Carlos Aliaga, Universidad de Zaragoza – I3A, and Desilico Labs, [carlos.aliaga@desilico.tech](mailto:carlos.aliaga@desilico.tech); Diego Gutierrez, Universidad de Zaragoza – I3A, [diegog@unizar.es](mailto:diegog@unizar.es).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART83 \$15.00

<https://doi.org/10.1145/3197517.3201282>

Additional Key Words and Phrases: Spatially-Correlated Transport, Non-Exponential Light Transport, Correlated Radiative Transfer

## ACM Reference Format:

Adrian Jarabo, Carlos Aliaga, and Diego Gutierrez. 2018. A Radiative Transfer Framework for Spatially-Correlated Materials. *ACM Trans. Graph.* 37, 4, Article 83 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201282>

## 1 INTRODUCTION

Volumetric appearances are ubiquitous in the real world, from translucent organic materials to clouds, smoke, or densely packed granular media. Voxel-based representations with anisotropic scattering functions [Heitz et al. 2015; Jakob et al. 2010] have been widely used in recent years to represent the appearance of complex geometries such as trees [Loubet and Neyret 2017; Neyret 1998], cloth and hair [Aliaga et al. 2017; Khungurn et al. 2015; Schröder et al. 2011; Zhao et al. 2011], or particles' aggregates [Meng et al. 2015; Moon et al. 2007; Müller et al. 2016].

Many translucent objects and participating media present a strong spatial correlation between scatterers<sup>1</sup> [Coquard and Baillis 2006; Knyazikhin et al. 1998; Lovejoy et al. 1995], where scatterers' densities are non-uniform within a differential volume. The aerosol of clouds, for instance, tends to form clusters, resulting in areas with very different optical thicknesses [Marshak et al. 1998]. As a result, the probability of a photon interacting with a scatterer inside each differential volume is also non-uniform, which in turn has a great effect in the final appearance, as Figure 1 shows.

Most previous works in graphics have assumed an uncorrelated distribution of scatterers, considering only spatial correlation at a

<sup>1</sup>Following other works' terminology, through the paper we use the term "scatterers" for all particles in the media, including perfect absorbers.



# Example-based Turbulence Style Transfer

SYUHEI SATO, DWANGO Co., Ltd., Dwango CG Research  
YOSHINORI DOBASHI, Hokkaido University and Dwango CG Research  
THEODORE KIM, Pixar Animation Studios  
TOMOYUKI NISHITA, Dwango CG Research and Hiroshima Shudo University

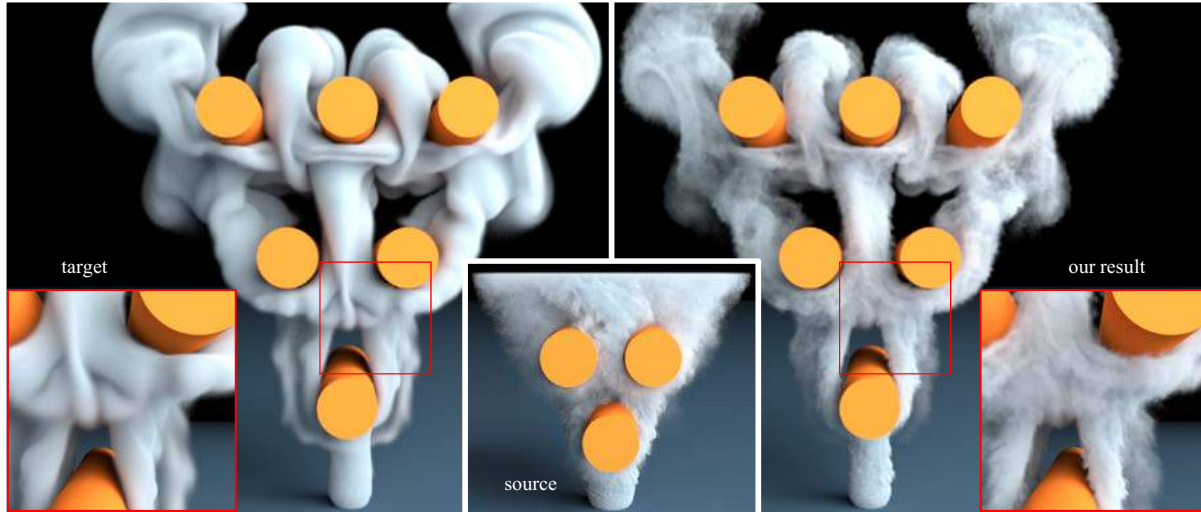


Fig. 1. Example of smoke interacting with obstacles. On the left is the original, low-resolution smoke (“target”). On the right, we transferred high-resolution turbulent motion from the middle image onto the low-resolution original. Closeups of the red square regions are shown in the insets.

Generating realistic fluid simulations remains computationally expensive, and animators can expend enormous effort trying to achieve a desired motion. To reduce such costs, several methods have been developed in which high-resolution turbulence is synthesized as a post process. Since global motion can then be obtained using a fast, low-resolution simulation, less effort is needed to create a realistic animation with the desired behavior. While much research has focused on accelerating the low-resolution simulation, the problem controlling the behavior of the turbulent, high-resolution motion has received little attention. In this paper, we show that *style transfer* methods from image editing can be adapted to transfer the turbulent style of an existing fluid simulation onto a new one. We do this by extending example-based image synthesis methods to handle velocity fields using a combination of patch-based and optimization-based texture synthesis. This approach allows us to take into account the incompressibility condition, which we

have found to be an important factor during synthesis. Using our method, a user can easily and intuitively create high-resolution fluid animations that have a desired turbulent motion.

CCS Concepts: • **Computing methodologies** → **Animation**; *Physical simulation*;

Additional Key Words and Phrases: fluid simulation, reusing existing fluid animations, incompressibility, texture synthesis, patch-based synthesis

## ACM Reference Format:

Syuhei Sato, Yoshinori Dobashi, Theodore Kim, and Tomoyuki Nishita. 2018. Example-based Turbulence Style Transfer. *ACM Trans. Graph.* 37, 4, Article 84 (August 2018), 9 pages. <https://doi.org/10.1145/3197517.3201398>

## 1 INTRODUCTION

Physically-based fluid simulation is used extensively in many production environments, such as movies. However, they remain computationally expensive, and animators must tediously repeat simulations multiple times in order to find parameter settings that produce a desired motion. In order to accelerate these design iterations, post-processing approaches have been proposed that add plausible turbulence [Kim et al. 2008; Narain et al. 2008; Schechter and Bridson 2008] and utilize guide-based formulations [Nielsen and Christensen 2010; Nielsen et al. 2009].

These methods all follow the same workflow: the overall motion is first efficiently authored at a low-resolution, and small-scale, high-frequency details are then synthesized as a post-process. While

Authors' addresses: Syuhei Sato, DWANGO Co., Ltd., Dwango CG Research, syuhei\_sato@dwango.co.jp; Yoshinori Dobashi, Hokkaido University, Dwango CG Research, doba@ime.ist.hokudai.ac.jp; Theodore Kim, Pixar Animation Studios, tkim@pixar.com; Tomoyuki Nishita, Dwango CG Research, Hiroshima Shudo University, nishita@shudo-u.ac.jp.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART84 \$15.00

<https://doi.org/10.1145/3197517.3201398>

# An Advection-Reflection Solver for Detail-Preserving Fluid Simulation

JONAS ZEHNDER, Université de Montréal

RAHUL NARAIN, University of Minnesota and Indian Institute of Technology Delhi

BERNHARD THOMASZEWSKI, Université de Montréal

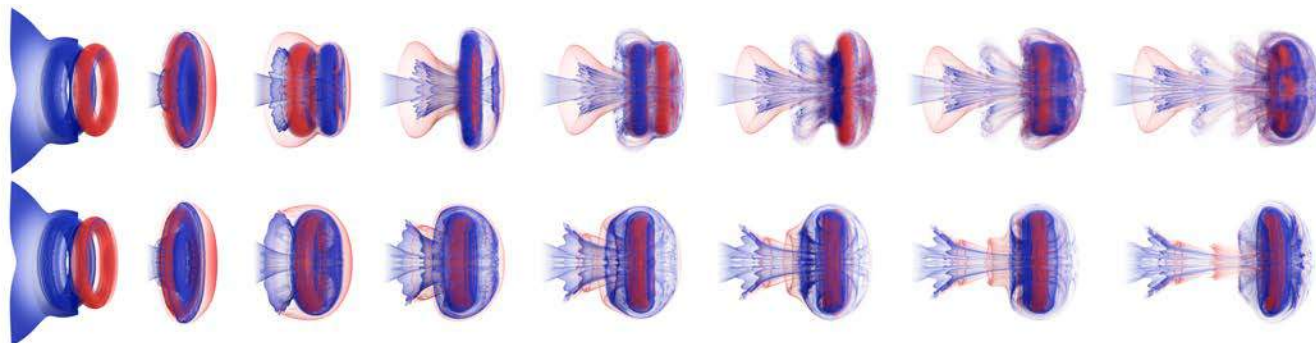


Fig. 1. Our new reflection solver applied to a vortex leap-frogging problem (*top row*). During 10s of simulation the two vortex rings move through each other multiple times and stay well separated. By contrast, in a standard advection-projection method with MacCormack advection (*bottom row*), the two vortices merge immediately and never separate afterwards.

Advection-projection methods for fluid animation are widely appreciated for their stability and efficiency. However, the projection step dissipates energy from the system, leading to artificial viscosity and suppression of small-scale details. We propose an alternative approach for detail-preserving fluid animation that is surprisingly simple and effective. We replace the energy-dissipating projection operator applied at the end of a simulation step by an *energy-preserving reflection* operator applied at mid-step. We show that doing so leads to two orders of magnitude reduction in energy loss, which in turn yields vastly improved detail-preservation. We evaluate our reflection solver on a set of 2D and 3D numerical experiments and show that it compares favorably to state-of-the-art methods. Finally, our method integrates seamlessly with existing projection-advection solvers and requires very little additional implementation.

CCS Concepts: • **Computing methodologies** → **Physical simulation**; **Computer graphics**; **Animation**;

Additional Key Words and Phrases: fluid simulation, advection, reflection, energy conservation

## ACM Reference Format:

Jonas Zehnder, Rahul Narain, and Bernhard Thomaszewski. 2018. An Advection-Reflection Solver for Detail-Preserving Fluid Simulation. *ACM Trans.*

Authors' addresses: Jonas Zehnder, Université de Montréal, jonas.zehnder@umontreal.ca; Rahul Narain, University of Minnesota, Indian Institute of Technology Delhi, narain@cse.iitd.ac.in; Bernhard Thomaszewski, Université de Montréal, bernhard@iro.umontreal.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://permissions.acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART85 \$15.00

<https://doi.org/10.1145/3197517.3201324>

*Graph.* 37, 4, Article 85 (August 2018), 8 pages. <https://doi.org/10.1145/3197517.3201324>

## 1 INTRODUCTION

Advection-projection methods are widely used for fluid animations in computer graphics. Splitting mass transport and conservation into different steps allows for stable and efficient integration, and advances in higher-order advection schemes (e.g. [Selle et al. 2008]) have greatly reduced the well-known numerical diffusion caused by the semi-Lagrangian advection step. However, the splitting of the time integration scheme itself induces numerical dissipation, as kinetic energy is transferred to divergent modes during advection and then lost after projection. This numerical dissipation manifests as rapid decay of large vortices and leads to suppression of small-scale swirling motion. Since visual complexity is a central goal in fluid animation, much effort has been spent on combating numerical dissipation: apart from higher-order advection schemes mentioned above, energy-preserving integration [Mullen et al. 2009], a posteriori correction of the velocity field [Fedkiw et al. 2001; Zhang et al. 2015], and injection of procedurally-generated detail [Kim et al. 2008b] are among the strategies that have been pursued so far.

In this work, we propose an alternative approach to detail-preserving fluid animation that is surprisingly simple and effective: we replace the energy-dissipating projection operator applied at the end of a simulation step by an *energy-preserving reflection* operator applied at mid-step; see Fig 2. We show that doing so leads to an order of magnitude reduction in divergent kinetic energy, which in turn leads to vastly improved preservation of vortices and small-scale detail. This analysis also exposes our method to be first-order structurally symmetric, which motivates an analogy to (and comparison with) symmetric projection methods for structure-preserving integration of constrained mechanical systems [Hairer et al. 2006].

# An Extended Partitioned Method for Conservative Solid-Fluid Coupling

MUZAFFER AKBAY, University of California, Riverside, USA

NICHOLAS NOBLES, University of California, Riverside, USA

VICTOR ZORDAN, Clemson University, USA

TAMAR SHINAR, University of California, Riverside, USA



Fig. 1. Left: Rigid balls of varying densities are dropped into a pool of water, achieving correct buoyancy. Center: A thin, elastic balloon is stressed and filled by a fast stream of water. Right: Pressure applied to a small piston of a hydraulic press is transmitted undiminished to a large piston through a fully enclosed and incompressible medium.

We present a novel extended partitioned method for two-way solid-fluid coupling, where the fluid and solid solvers are treated as black boxes with limited exposed interfaces, facilitating modularity and code reusability. Our method achieves improved stability and extended range of applicability over standard partitioned approaches through three techniques. First, we couple the black-box solvers through a small, reduced-order monolithic system, which is constructed on the fly from input/output pairs generated by the solid and fluid solvers. Second, we use a conservative, impulse-based interaction term to couple the solid and fluid rather than typical pressure-based forces. We show that both of these techniques significantly improve stability and reduce the number of iterations needed for convergence. Finally, we propose a novel boundary pressure projection method that allows for the partitioned simulation of a fully enclosed fluid coupled to a dynamic solid, a scenario that has been problematic for partitioned methods. We demonstrate the benefits of our extended partitioned method by coupling Eulerian fluid solvers for smoke and water to Lagrangian solid solvers for volumetric and thin deformable and rigid objects in a variety of challenging scenarios. We further demonstrate our method by coupling a Lagrangian SPH fluid solver to a rigid body solver.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: solid-fluid coupling, partitioned simulation, reduced-order model

Authors' addresses: Muzaffer Akbay, University of California, Riverside, 900 University Ave. Riverside, California, 92521, USA, akbaym@cs.ucr.edu; Nicholas Nobles, University of California, Riverside, 900 University Ave. Riverside, California, 92521, USA, nnobl002@ucr.edu; Victor Zordan, Clemson University, Clemson University, Clemson, South Carolina, 29634, USA, vbz@g.clemson.edu; Tamar Shinar, University of California, Riverside, 900 University Ave. Riverside, California, 92521, USA, shinar@cs.ucr.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2018/8-ART86 \$15.00

<https://doi.org/10.1145/3197517.3201345>

## ACM Reference Format:

Muzaffer Akbay, Nicholas Nobles, Victor Zordan, and Tamar Shinar. 2018. An Extended Partitioned Method for Conservative Solid-Fluid Coupling. *ACM Trans. Graph.* 37, 4, Article 86 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201345>

## 1 INTRODUCTION

Physically-based simulation of coupled fluids and solids has been an area of active research in recent years, both in computer graphics and computational physics, and many solution approaches have been investigated [Hou et al. 2012]. The techniques generally fall into two categories, monolithic and partitioned approaches. In monolithic approaches, equations for the solid and fluid are combined into one system and solved simultaneously, implicitly coupling internal solid and fluid forces to the interaction forces between them. On the other hand, partitioned approaches employ separate solvers for the solid and fluid and interleave calls to each solver to achieve either a weak coupling, or if iterated to convergence, a strong coupling of the solvers.

Both the monolithic and partitioned approaches have advantages and disadvantages in the simulation of two-way solid-fluid coupling. Monolithic approaches have the advantage of better stability, as stiff coupling terms can be resolved simultaneously with other forces in the system, but they also incur significant cost in terms of development and do not fully leverage existing codes. On the other hand, partitioned approaches have the advantage of being able to reuse previously developed fluid and solid solvers in a black-box fashion as long as they expose appropriate interfaces. However, they suffer from poorer stability properties, and in challenging scenarios, they may require many iterations per time step or worse, fail to converge [Banks et al. 2014]. When strong coupling is desired, a large number of iterations may be required for convergence, negatively impacting performance. Despite their drawbacks, partitioned approaches remain an important class of methods due to the appeal of reusing existing software, and recent work on partitioned



# Scalable Laplacian Eigenfluids

QIAODONG CUI, University of California, Santa Barbara

PRADEEP SEN, University of California, Santa Barbara

THEODORE KIM, Pixar Animation Studios

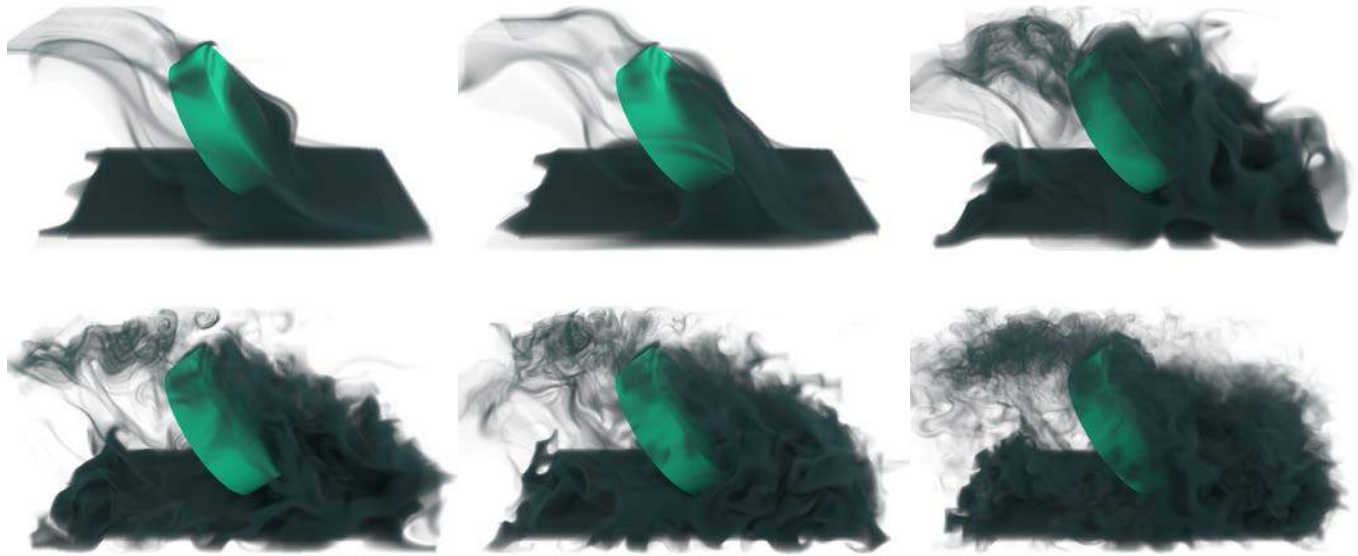


Fig. 1. PADDLE WHEEL scene: Top row:  $r = 100$ ,  $r = 200$ ,  $r = 1000$  basis functions are used. Bottom row:  $r = 2000$ ,  $r = 4000$ ,  $r = 12600$  basis functions are used. Previous approaches have only been able to achieve  $r \approx 500$ , and would have needed 2.25 TB of memory to simulate the  $r = 12600$  case.

The Laplacian Eigenfunction method for fluid simulation, which we refer to as *Eigenfluids*, introduced an elegant new way to capture intricate fluid flows with near-zero viscosity. However, the approach does not scale well, as the memory cost grows prohibitively with the number of eigenfunctions. The method also lacks generality, because the dynamics are constrained to a closed box with Dirichlet boundaries, while open, Neumann boundaries are also needed in most practical scenarios. To address these limitations, we present a set of analytic eigenfunctions that supports uniform Neumann and Dirichlet conditions along each domain boundary, and show that by carefully applying the discrete sine and cosine transforms, the storage costs of the eigenfunctions can be made completely negligible. The resulting algorithm is both faster and more memory-efficient than previous approaches, and able to achieve lower viscosities than similar pseudo-spectral methods. We are able to surpass the scalability of the original Laplacian Eigenfunction

approach by over two orders of magnitude when simulating rectangular domains. Finally, we show that the formulation allows forward scattering to be directed in a way that is not possible with any other method.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: fluid simulation, physically based animation

## ACM Reference Format:

Qiaodong Cui, Pradeep Sen, and Theodore Kim. 2018. Scalable Laplacian Eigenfluids. *ACM Trans. Graph.* 37, 4, Article 87 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201352>

## 1 INTRODUCTION

While much progress has been made in fluid simulation over the last decade, efficiently simulating inviscid fluids that do not suffer from excessive numerical dissipation remains an ongoing challenge. Even in simple scenarios, the results can look unconvincingly viscous if any extraneous numerical diffusion is present.

Methods based on the traditional method of Stam [1999] usually compensate for the dissipated energy using techniques such as vorticity confinement [Fedkiw et al. 2001] or IVOCK [Zhang et al. 2015]. Alternatively, methods can be devised that are non-dissipative by construction, though these can involve asymmetric linear solves and non-linear Newton iterations [Mullen et al. 2009].

Authors' addresses: Qiaodong Cui, University of California, Santa Barbara, [qiaodong@umail.ucsb.edu](mailto:qiaodong@umail.ucsb.edu); Pradeep Sen, University of California, Santa Barbara, [psen@ece.ucsb.edu](mailto:psen@ece.ucsb.edu); Theodore Kim, Pixar Animation Studios, [tkim@pixar.com](mailto:tkim@pixar.com).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART87 \$15.00

<https://doi.org/10.1145/3197517.3201352>

# Creating and Chaining Camera Moves for Quadrotor Videography

KE XIE, Shenzhen University

HAO YANG, Shenzhen University

SHENGQIU HUANG, Shenzhen University

DANI LISCHINSKI, The Hebrew University of Jerusalem

MARC CHRISTIE, Univ Rennes, INRIA, CNRS

KAI XU, Shenzhen University and National University of Defense Technology

MINGLUN GONG, Memorial University of Newfoundland

DANIEL COHEN-OR, Shenzhen University and Tel Aviv University

HUI HUANG\*, Shenzhen University

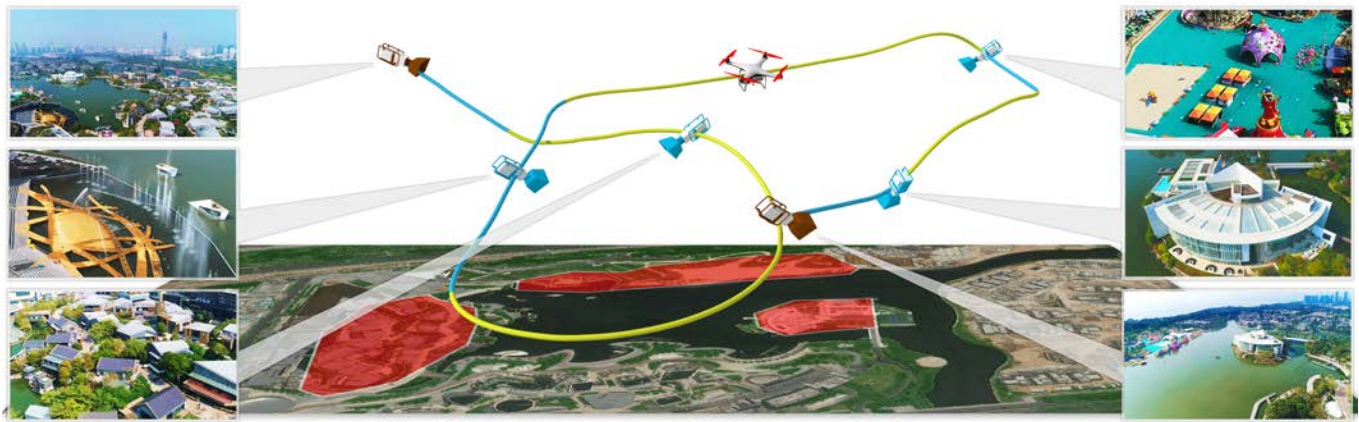


Fig. 1. Given several landmarks (red areas on the satellite image) and a pair of start and end views (brown camera icons), we generate a suitable camera move (in yellow) for capturing each landmark, and optimally connect them into a continuous and smooth path using transition trajectories (in blue).

Capturing aerial videos with a quadrotor-mounted camera is a challenging creative task, as it requires the simultaneous control of the quadrotor's motion and the mounted camera's orientation. Letting the drone follow a pre-planned trajectory is a much more appealing option, and recent research has proposed a number of tools designed to automate the generation of feasible camera motion plans; however, these tools typically require the user to specify and edit the camera path, for example by providing a complete and ordered sequence of key viewpoints.

\*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Ke Xie, Shenzhen University; Hao Yang, Shenzhen University; Shengqiu Huang, Shenzhen University; Dani Lischinski, The Hebrew University of Jerusalem; Marc Christie, Univ Rennes, INRIA, CNRS; Kai Xu, Shenzhen University, National University of Defense Technology; Minglun Gong, Memorial University of Newfoundland; Daniel Cohen-Or, Shenzhen University, Tel Aviv University; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART88 \$15.00

<https://doi.org/10.1145/3197517.3201286>

In this paper, we propose a higher level tool designed to enable even novice users to easily capture compelling aerial videos of large-scale outdoor scenes. Using a coarse 2.5D model of a scene, the user is only expected to specify starting and ending viewpoints and designate a set of landmarks, with or without a particular order. Our system automatically generates a diverse set of candidate local *camera moves* for observing each landmark, which are collision-free, smooth, and adapted to the shape of the landmark. These moves are guided by a landmark-centric view quality field, which combines visual interest and frame composition. An optimal global camera trajectory is then constructed that chains together a sequence of local camera moves, by choosing one move for each landmark and connecting them with suitable transition trajectories. This task is formulated and solved as an instance of the Set Traveling Salesman Problem.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; Interest point and salient region detections; Graphics systems and interfaces;

Additional Key Words and Phrases: Quadrotor videography, trajectory planning, camera control, scene navigation, visual exploration

## ACM Reference Format:

Ke Xie, Hao Yang, Shengqiu Huang, Dani Lischinski, Marc Christie, Kai Xu, Minglun Gong, Daniel Cohen-Or, and Hui Huang. 2018. Creating and Chaining Camera Moves for Quadrotor Videography. *ACM Trans. Graph.* 37, 4, Article 88 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201286>

# Learning Three-Dimensional Flow for Interactive Aerodynamic Design

NOBUYUKI UMETANI, Autodesk Research  
BERND BICKEL, IST Austria

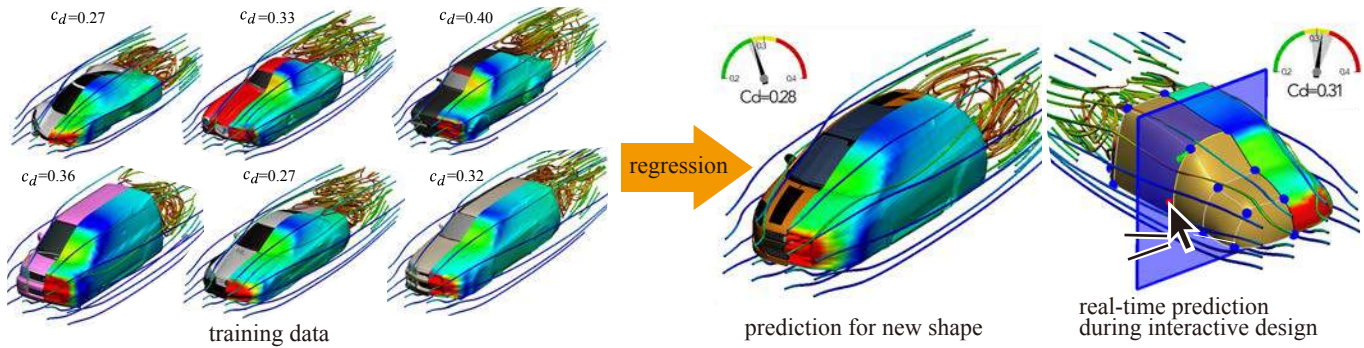


Fig. 1. Our interactive design system allows users to interactively design and optimize a free-form 3D shape while the pressure on the surface (color-coded) and the velocity field (illustrated as stream lines) are accurately predicted in real-time. Our data-driven model is based on a novel representation for 3D shapes and can be robustly trained from a set of exemplars.

We present a data-driven technique to instantly predict how fluid flows around various three-dimensional objects. Such simulation is useful for computational fabrication and engineering, but is usually computationally expensive since it requires solving the Navier-Stokes equation for many time steps. To accelerate the process, we propose a machine learning framework which predicts aerodynamic forces and velocity and pressure fields given a three-dimensional shape input. Handling detailed free-form three-dimensional shapes in a data-driven framework is challenging because machine learning approaches usually require a consistent parametrization of input and output. We present a novel PolyCube maps-based parametrization that can be computed for three-dimensional shapes at interactive rates. This allows us to efficiently learn the nonlinear response of the flow using a Gaussian process regression. We demonstrate the effectiveness of our approach for the interactive design and optimization of a car body.

CCS Concepts: • **Computing methodologies** → **Gaussian processes**; *Real-time simulation*; *Physical simulation*; • **Applied computing** → *Engineering*;

Additional Key Words and Phrases: machine learning, fluid simulation, Gaussian process, parameterization

## ACM Reference Format:

Nobuyuki Umetani and Bernd Bickel. 2018. Learning Three-Dimensional Flow for Interactive Aerodynamic Design. *ACM Trans. Graph.* 37, 4, Article 89 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201325>

Authors' addresses: Nobuyuki Umetani, Autodesk Research, Toronto, Canada, n.umatani@gmail.com; Bernd Bickel, IST Austria, Vienna, Austria, bernd.bickel@ist.ac.at.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART89 \$15.00

<https://doi.org/10.1145/3197517.3201325>

## 1 INTRODUCTION

Computational Fluid Dynamics (CFD) has become an indispensable component in the field of engineering and computational fabrication for design evaluation and optimization. Successful applications include improving the aerodynamics of cars and airplanes, refining heating, ventilation and air conditioning in architectural planning, or the optimization of casting and injection molding in manufacturing. Many of these applications would benefit from an interactive design approach. However, analyzing a design is usually computationally expensive, as it requires repetitively solving a system with many degrees of freedom over many time steps. Although modern design tools allow the user to edit the shape interactively, computation times often prohibit the interactive visualization, exploration, and optimization of shapes and their surroundings based on CFD simulations. This problem is extremely challenging, even when real-time simulation capabilities are available, i.e., the simulation of individual time-steps at real-time rates, because evaluating a design in simulation may require many time steps.

In this paper, we suggest a novel approach to achieve real-time CFD performance based on a *data-driven* method. While CFD generally refers to fluid simulation for engineering purposes, we limit the scope to the evaluation of aerodynamic forces and flow around an object given that it is moving at a constant speed in a single-phase fluid (e.g., air flow around cars or airplanes). Our approach enables drastically faster evaluation by constructing a reduced model between input and output based on pre-existing data. Our model takes a 3D shape as the input, and outputs a drag coefficient, fluid pressure on the surface and velocity field around the shape. The nonlinearly and globally coupled relationship between input and output is learned using the Gaussian Process (GP).

A key challenge in applying machine learning to CFD problems is the parameterization of the input and output. As most machine learning approaches take fixed dimensional vectors as input and



# Optimizing for Aesthetically Pleasing Quadrotor Camera Motion

CHRISTOPH GEBHARDT, STEFAN STEVŠIĆ, and OTMAR HILLIGES, AIT Lab, ETH Zürich

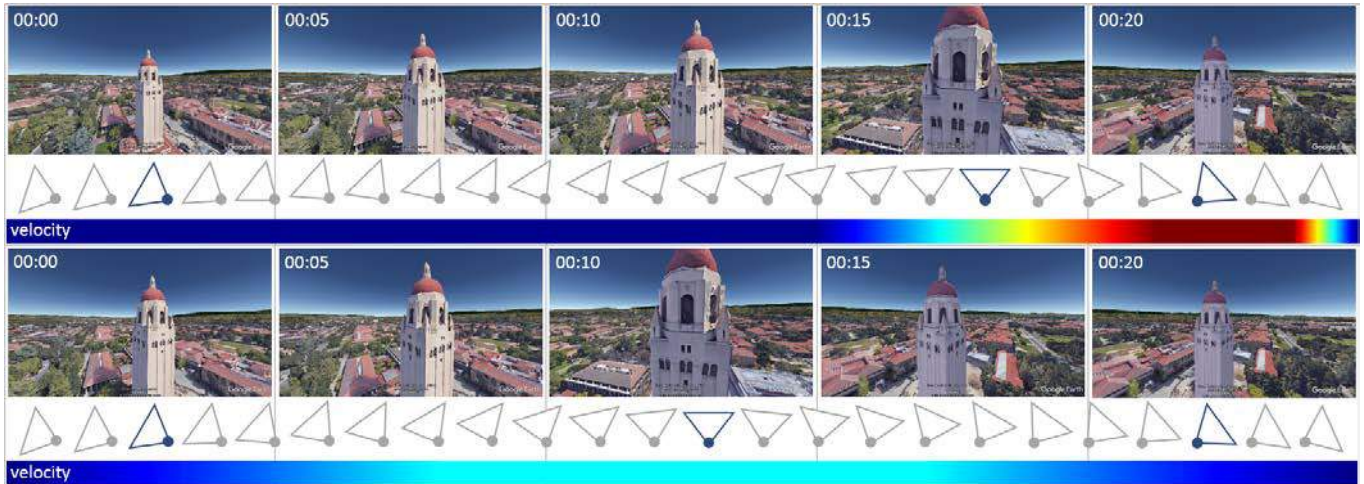


Fig. 1. Quadrotor camera tools generate trajectories based on user-specified keyframes in time and space. Reasoning about spatio-temporal distances is hard for users and can lead to visually unappealing results and fluctuating camera velocities. *Top row*: user-specified keyframes (blue) are positioned in time, such that the camera first moves too slow and then needs to accelerate drastically to reach the specified end-point. *Bottom row*: results of our method which automatically positions keyframes (blue) in time such that the camera moves smoothly over the entire trajectory (illustrative example).

In this paper we first contribute a large scale online study ( $N \approx 400$ ) to better understand aesthetic perception of aerial video. The results indicate that it is paramount to optimize smoothness of trajectories across all keyframes. However, for experts timing control remains an essential tool. Satisfying this dual goal is technically challenging because it requires giving up desirable properties in the optimization formulation. Second, informed by this study we propose a method that optimizes positional and temporal reference fit *jointly*. This allows to generate globally smooth trajectories, while retaining user control over reference timings. The formulation is posed as a variable, infinite horizon, contour-following algorithm. Finally, a comparative lab study indicates that our optimization scheme outperforms the state-of-the-art in terms of perceived usability and preference of resulting videos. For novices our method produces smoother and better looking results and also experts benefit from generated timings.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Motion path planning**; **Robotic planning**; • **Computer systems organization** → External interfaces for robotics;

Additional Key Words and Phrases: computational design, aerial videography, quadrotor camera tools, trajectory optimization

Authors' address: Christoph Gebhardt, cgebhardt@inf.ethz.ch; Stefan Stevšić, stevsics@inf.ethz.ch; Otmar Hilliges, hilliges@inf.ethz.ch, AIT Lab, ETH Zürich.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.  
0730-0301/2018/8-ART90 \$15.00  
<https://doi.org/10.1145/3197517.3201390>

## ACM Reference Format:

Christoph Gebhardt, Stefan Stevšić, and Otmar Hilliges. 2018. Optimizing for Aesthetically Pleasing Quadrotor Camera Motion. *ACM Trans. Graph.* 37, 4, Article 90 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201390>

## 1 INTRODUCTION

Camera quadrotors have become a mainstream technology but fine-grained control of such camera drones for aerial videography is a high-dimensional and hence difficult task. In response several tools have been proposed to plan quadrotor shots by defining keyframes in virtual environments [Gebhardt et al. 2016; Gebhardt and Hilliges 2018; Joubert et al. 2015; Roberts and Hanrahan 2016]. This input is then used in an optimization algorithm to automatically generate quadrotor and camera trajectories. Intuitively, smooth camera motion is an obvious factor impacting the visual quality of a shot. This intuition alongside expert-feedback [Joubert et al. 2015] and literature on (aerial) cinematography [Arijon 1976; Audronis 2014; Hennessy 2015] forms the basis for most existing quadrotor tools. These take a spline representation, connecting user specified keyframes, and optimize higher derivatives of these splines, such as jerk.

Balasubramanian et al. [2015] define global smoothness as “a quality related to the continuity or non-intermittency of a movement, independent of its amplitude and duration”. However, because keyframe timings are kept fixed in current quadrotor camera optimization schemes [Gebhardt et al. 2016; Joubert et al. 2015], or close to the user input [Roberts and Hanrahan 2016], smooth motion can only be generated subject to these hard-constraints. This can cause strong variation of camera velocities across different trajectory segments and result in visually unpleasant videos.

# Integer-Only Cross Field Computation

NAHUM FARCHI and MIRELA BEN-CHEN, Technion Institute of Technology

We propose a new iterative algorithm for computing smooth cross fields on triangle meshes that is simple, easily parallelizable on the GPU, and finds solutions with lower energy and fewer cone singularities than state-of-the-art methods. Our approach is based on a formal equivalence, which we prove, between two formulations of the optimization problem. This equivalence allows us to eliminate the real variables and design an efficient grid search algorithm for the cone singularities. We leverage a recent graph-theoretical approximation of the *resistance distance matrix* of the triangle mesh to speed up the computation and enable a trade-off between the computation time and the smoothness of the output.

CCS Concepts: • **Computing methodologies** → **Mesh models**;

Additional Key Words and Phrases: cross fields, singularities, quad remeshing, digital geometry processing, resistance distance

## ACM Reference Format:

Nahum Farchi and Mirela Ben-Chen, Technion Institute of Technology. 2018. Integer-Only Cross Field Computation. *ACM Trans. Graph.* 37, 4, Article 91 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201375>

## 1 INTRODUCTION

Directional fields, and especially *cross fields*, are important objects in geometry processing. They are used in many applications, from quadrangular remeshing to non-photorealistic rendering [Vaxman et al. 2016]. Computing smooth cross fields on triangle meshes is challenging, as the problem formulation inherently depends on *integer* variables to encode the invariance of the crosses to rotations by integer multiples of  $\pi/2$ .

A popular approach, suggested by Bommes et al. [2009] (MIQ), formulates a mixed-integer optimization problem and solves it greedily to compute the cross field. While highly efficient and effective, the greedy solution can lead to sub-optimal results, as in Fig. 1 (top). Alternatively, Crane et al. [2010] (TCODS) [Crane et al. 2010] posed the problem in terms of *angle defects* due to parallel transport on closed cycles, leading to a sparse linear least squares problem that is solved efficiently when the defects are known.

We show that if the angle defects are *unknown*, and there are no directional constraints, these two optimization problems are *equivalent*. Furthermore, by eliminating the real variables, we remain with an *integer only* optimization problem. We use this insight to design a new iterative algorithm for minimizing the energy that is simple, easily parallelizable on the GPU, and finds solutions with lower energy and fewer singularities than MIQ, e.g. Fig. 1 (bottom). Finally, we show the connection of the minimized energy to the *resistance distance matrix* of the triangle mesh, and leverage a recent

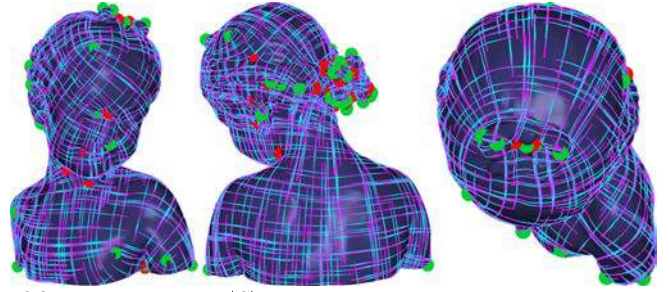
graph theoretical approximation to speed up the computation and allow us to trade-off the computation time and the quality of the resulting cross field.

### 1.1 Related Work

Cross field computation, and directional field computation in general, has seen a surge of research in recent years. A recent review [Vaxman et al. 2016] covers the latest developments, and we therefore focus our literature review on methods closest to our approach.

*Angle based representation.* A popular formulation of the cross field computation problem is to represent every cross as an *angle* with respect to a fixed local orthogonal frame. Since crosses are invariant to rotations by integer multiples of  $\pi/2$ , such a representation has an inherent phase ambiguity. Therefore, finding a *smooth* assignment of crosses inevitably requires taking into account these unknown integer phases, leading to optimization problems with integer variables. Bommes et al. [2009] suggested one of the first efficient methods to tackle these optimization problems in the context of cross field generation, by greedily rounding to an integer one variable per iteration and resolving the system. Our approach optimizes the same energy greedily, albeit using a different algorithm that guarantees that there exists no modification of a single  $\pm\pi/2$  singularity's position that reduces the energy. This leads to lower energy values and better singularity placement. A different angle

MIQ,  $E = 82.46$ ,  $|S| = 86$



IOQe  $\epsilon = .5$ ,  $E = 59.10$ ,  $|S| = 28$

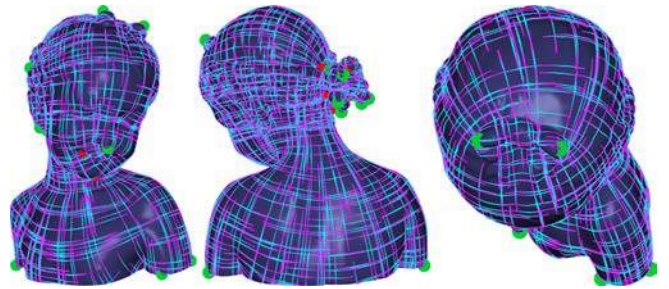


Fig. 1. Our iterative optimization (bottom) finds a solution with lower energy, and fewer singularities than MIQ (top).

Author's address: Nahum Farchi and Mirela Ben-Chen, Technion Institute of Technology.



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART91  
<https://doi.org/10.1145/3197517.3201375>



# Quadrangulation through Morse-Parameterization Hybridization

XIANZHONG FANG, State Key Lab of CAD&CG, Zhejiang University

HUJUN BAO\*, State Key Lab of CAD&CG, Zhejiang University

YIYING TONG, Michigan State University

MATHIEU DESBRUN, Caltech

JIN HUANG\*, State Key Lab of CAD&CG, Zhejiang University

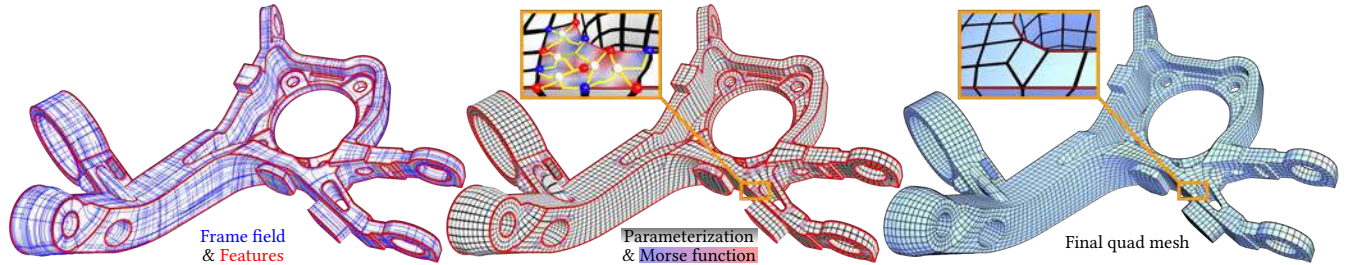


Fig. 1. **Fast and robust hybrid quadrangulation via periodic vector field.** We present an approach for efficiently producing frame- and feature-aligned quadrilateral meshes from input triangle meshes. By combining the efficiency of parameterization-based methods and the theoretical guarantees of Morse-based methods, our hybrid method robustly and efficiently extracts a non-degenerate quad mesh from a 4D periodic vector field computed from the input geometry and its frame field (here, a mechanical part model with complex feature lines).

We introduce an approach to quadrilateral meshing of arbitrary triangulated surfaces that combines the theoretical guarantees of Morse-based approaches with the practical advantages of parameterization methods. We first construct, through an eigensolver followed by a few Gauss-Newton iterations, a periodic four-dimensional vector field that aligns with a user-provided frame field and/or a set of features over the input mesh. A field-aligned parameterization is then greedily computed along a spanning tree based on the Dirichlet energy of the optimal periodic vector field, from which quad elements are efficiently extracted over most of the surface. The few regions not yet covered by elements are then upsampled and the first component of the periodic vector field is used as a Morse function to extract the remaining quadrangles. This hybrid parameterization- and Morse-based quad meshing method is not only fast (the parameterization is greedily constructed, and the Morse function only needs to be upsampled in the few uncovered patches), but is guaranteed to provide a feature-aligned quad mesh with non-degenerate cells that closely matches the input frame field over an arbitrary surface. We show that our approach is much faster than Morse-based techniques since it does not require a densely tessellated input mesh, and is significantly more robust than parameterization-based techniques on models with complex features.

CCS Concepts: • **Mathematics of computing** → **Mesh generation**;

Additional Key Words and Phrases: Quadrangulation, periodic vector fields, frame fields, parameterization, Morse-Smale complex.

\*Corresponding authors: {bao,hj}@cad.zju.edu.cn.

Authors' addresses: X. Fang, H. Bao, J. Huang, State Key Lab of CAD&CG, Zhejiang University, China; Y. Tong, Michigan State University, USA; M. Desbrun, Caltech, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART92 \$15.00

<https://doi.org/10.1145/3197517.3201354>

## ACM Reference Format:

Xianzhong Fang, Hujun Bao\*, Yiying Tong, Mathieu Desbrun, and Jin Huang\*. 2018. Quadrangulation through Morse-Parameterization Hybridization. *ACM Trans. Graph.* 37, 4, Article 92 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201354>

## 1 INTRODUCTION

To this day, fully automated quadrangulation of surfaces with features and arbitrary guiding frame fields remains a challenge: producing a pure quad mesh with non-degenerate cells quickly and reliably while tightly adhering to a prescribed frame field indicating the desired edge directions and lengths is too often impossible without user interaction with current methods.

A first family of techniques (that we refer to as *parameterization-based quad meshing*) relies on the singularity structure of the input frame field to derive a field-aligned parameterization, typically involving mixed integer optimization [Bommes et al. 2013a, 2009; Kälberer et al. 2007], from which a quad mesh can be extracted through isocontouring. While this general approach can be made computationally efficient, non-degeneracy of the parameterization is never guaranteed, in particular due to possible inconsistency between feature/boundary alignment, holonomy conditions and user constraints. A competing family of approaches (that we will denote as *Morse-based quad meshing* [Dong et al. 2006; Huang et al. 2008; Ling et al. 2014; Zhang et al. 2010]) exploits instead a key property of any Morse function: its induced Morse-Smale Complex (MSC) always corresponds to a valid quadrangulation [Edelsbrunner et al. 2003]—and so does its quasi-dual MSC [Dong et al. 2006]—as long as the function assumes different values for adjacent vertices, which can always be enforced through symbolic perturbation [Zomorodian 2009]. However, Morse-based methods require a very densely tessellated initial surface to be able to capture the highest frequencies of a Morse function (see Fig.2). Consequently, one can only



# Singularity-Constrained Octahedral Fields for Hexahedral Meshing

HENG LIU, RWTH Aachen University  
PAUL ZHANG, MIT  
EDWARD CHIEN, MIT  
JUSTIN SOLOMON, MIT  
DAVID BOMMES, RWTH Aachen University

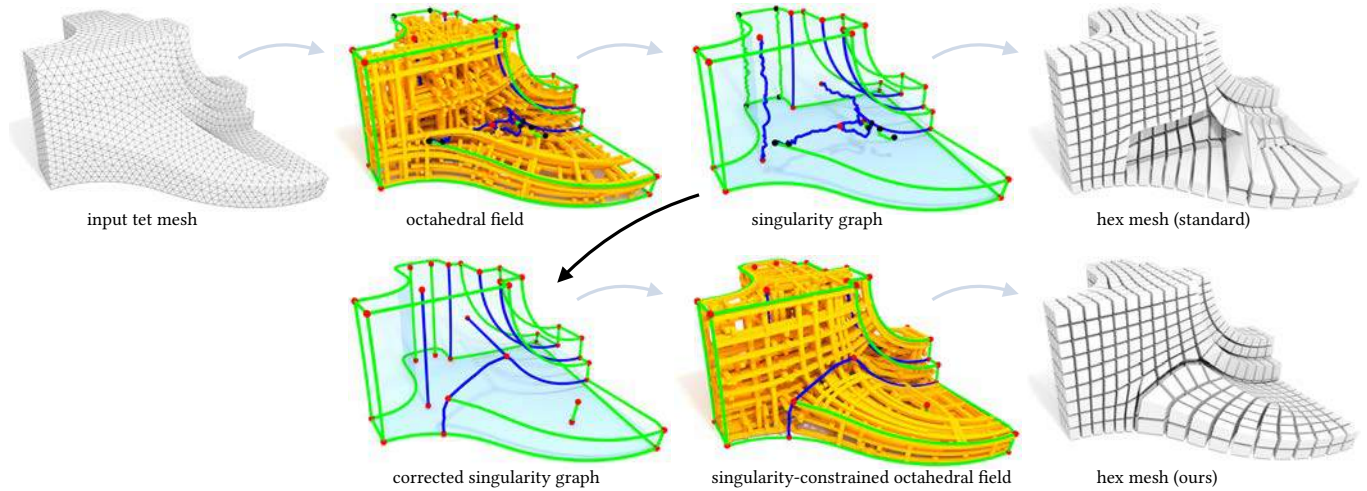


Fig. 1. Overview. Octahedral fields often exhibit singularities that are invalid for hex meshing, inducing non-hex elements and poor quality (top row). Our algorithm (bottom row) starts from a corrected singularity graph and generates a hex-meshable oct. field, resulting in a valid and less distorted hex mesh.

Despite high practical demand, algorithmic hexahedral meshing with guarantees on robustness and quality remains unsolved. A promising direction follows the idea of integer-grid maps, which pull back the Cartesian hexahedral grid formed by integer isoplanes from a parametric domain to a surface-conforming hexahedral mesh of the input object. Since directly optimizing for a high-quality integer-grid map is mathematically challenging, the construction is usually split into two steps: (1) generation of a surface-aligned octahedral field and (2) generation of an integer-grid map that best aligns to the octahedral field. The main robustness issue stems from the fact that smooth octahedral fields frequently exhibit singularity graphs that are not appropriate for hexahedral meshing and induce heavily degenerate integer-grid maps. The first contribution of this work is an enumeration of all local configurations that exist in hex meshes with bounded edge valence, and a generalization of the Hopf-Poincaré formula to octahedral fields, leading to necessary local and global conditions for the hex-meshability

of an octahedral field in terms of its singularity graph. The second contribution is a novel algorithm to generate octahedral fields with prescribed hex-meshable singularity graphs, which requires the solution of a large non-linear mixed-integer algebraic system. This algorithm is an important step toward robust automatic hexahedral meshing since it enables the generation of a hex-meshable octahedral field.

CCS Concepts: • **Computing methodologies** → **Mesh models; Mesh geometry models; Volumetric models;**

Additional Key Words and Phrases: hexahedral meshing, octahedral fields, singularity graph, integer-grid maps

## ACM Reference Format:

Heng Liu, Paul Zhang, Edward Chien, Justin Solomon, and David Bommes. 2018. Singularity-Constrained Octahedral Fields for Hexahedral Meshing. *ACM Trans. Graph.* 37, 4, Article 93 (August 2018), 17 pages. <https://doi.org/10.1145/3197517.3201344>

## 1 INTRODUCTION

A key step in physical simulation and physically-based animation involves dividing the spatial domain of the problem into a mesh whose elements are associated to the unknown variables. This discretization procedure has a strong bearing on the stability and fidelity of the result: Poorly-shaped, incorrectly-oriented, or unevenly-sized elements all can lead to failures in the simulation process.

A subtle consideration in meshing involves the *type* of element. In particular, there are many structures into which the domain of

Authors' addresses: Heng Liu, RWTH Aachen University, [liu@ices.rwth-aachen.de](mailto:liu@ices.rwth-aachen.de); Paul Zhang, MIT, [pzp1@gmail.com](mailto:pzpzp1@gmail.com); Edward Chien, MIT, [edward.d.chien@gmail.com](mailto:edward.d.chien@gmail.com); Justin Solomon, MIT, [jsolomon@mit.edu](mailto:jsolomon@mit.edu); David Bommes, RWTH Aachen University, [bommes@cs.rwth-aachen.de](mailto:bommes@cs.rwth-aachen.de).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART93 \$15.00 <https://doi.org/10.1145/3197517.3201344>

# Water Surface Wavelets

STEFAN JESCHKE, NVIDIA

TOMÁŠ SKŘIVAN, IST Austria

MATTHIAS MÜLLER-FISCHER, NVIDIA

NUTTAPONG CHENTANEZ, NVIDIA

MILES MACKLIN, NVIDIA

CHRIS WOJTAN, IST Austria



Fig. 1. A frame from a real-time wave animation with the simulation grid overlaid on top. We compute the wave motions at a resolution several orders of magnitude finer than the simulation variables.

The current state of the art in real-time two-dimensional water wave simulation requires developers to choose between efficient Fourier-based methods, which lack interactions with moving obstacles, and finite-difference or finite element methods, which handle environmental interactions but are significantly more expensive. This paper attempts to bridge this long-standing gap between complexity and performance, by proposing a new wave simulation method that can faithfully simulate wave interactions with moving

Authors' addresses: Stefan Jeschke, NVIDIA, [jeschke@stefan-jeschke.com](mailto:jeschke@stefan-jeschke.com); Tomáš Skřivan, IST Austria, [tomas.skřivan@ist.ac.at](mailto:tomas.skřivan@ist.ac.at); Matthias Müller-Fischer, NVIDIA, [matthiasm@nvidia.com](mailto:matthiasm@nvidia.com); Nuttapong Chentanez, NVIDIA, [nchentanez@nvidia.com](mailto:nchentanez@nvidia.com); Miles Macklin, NVIDIA, [mmacklin@nvidia.com](mailto:mmacklin@nvidia.com); Chris Wojtan, IST Austria, [wojtan@ist.ac.at](mailto:wojtan@ist.ac.at).



This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART94  
<https://doi.org/10.1145/3197517.3201336>

obstacles in real time while simultaneously preserving minute details and accommodating very large simulation domains.

Previous methods for simulating 2D water waves directly compute the change in height of the water surface, a strategy which imposes limitations based on the CFL condition (fast moving waves require small time steps) and Nyquist's limit (small wave details require closely-spaced simulation variables). This paper proposes a novel wavelet transformation that discretizes the liquid motion in terms of amplitude-like functions that vary over *space*, *frequency*, and *direction*, effectively generalizing Fourier-based methods to handle local interactions. Because these new variables change much more slowly over space than the original water height function, our change of variables drastically reduces the limitations of the CFL condition and Nyquist limit, allowing us to simulate highly detailed water waves at very large visual resolutions. Our discretization is amenable to fast summation and easy to parallelize. We also present basic extensions like pre-computed wave paths and two-way solid fluid coupling. Finally, we argue that our discretization provides a convenient set of variables for artistic manipulation, which we illustrate with a novel wave-painting interface.

# tempoGAN: A Temporally Coherent, Volumetric GAN for Super-resolution Fluid Flow

YOU XIE\*, Technical University of Munich  
ERIK FRANZ\*, Technical University of Munich  
MENGYU CHU\*, Technical University of Munich  
NILS THUEREY, Technical University of Munich

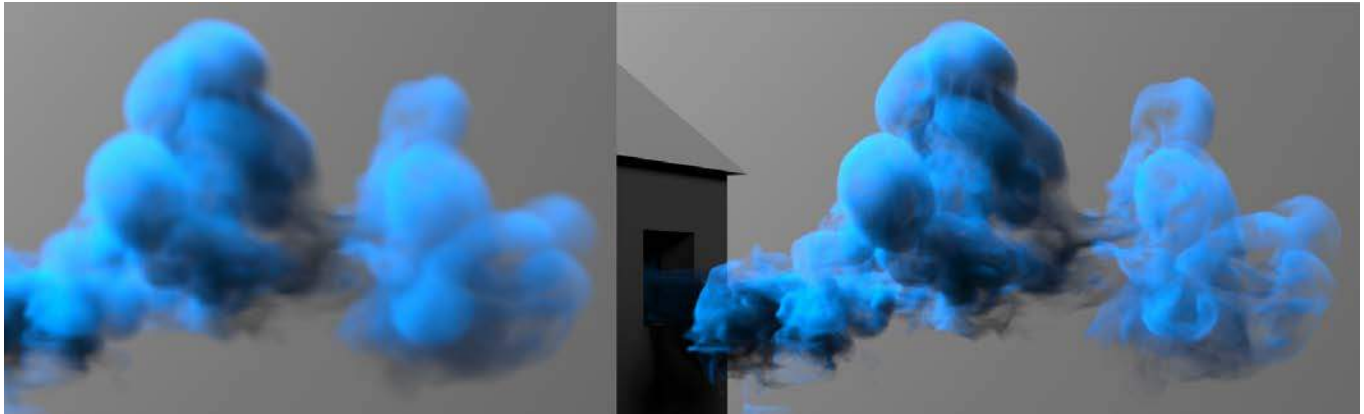


Fig. 1. Our convolutional neural network learns to generate highly detailed, and temporally coherent features based on a low-resolution field containing a single time-step of density and velocity data. We introduce a novel discriminator that ensures the synthesized details change smoothly over time.

We propose a temporally coherent generative model addressing the super-resolution problem for fluid flows. Our work represents a first approach to synthesize four-dimensional physics fields with neural networks. Based on a conditional generative adversarial network that is designed for the inference of three-dimensional volumetric data, our model generates consistent and detailed results by using a novel temporal discriminator, in addition to the commonly used spatial one. Our experiments show that the generator is able to infer more realistic high-resolution details by using additional physical quantities, such as low-resolution velocities or vorticities. Besides improvements in the training process and in the generated outputs, these inputs offer means for artistic control as well. We additionally employ a physics-aware data augmentation step, which is crucial to avoid overfitting and to reduce memory requirements. In this way, our network learns to generate advected quantities with highly detailed, realistic, and temporally coherent features. Our method works instantaneously, using only a single time-step of low-resolution fluid data. We demonstrate the abilities of our

method using a variety of complex inputs and applications in two and three dimensions.

CCS Concepts: • **Computing methodologies** → **Neural networks; Physical simulation**;

Additional Key Words and Phrases: physics-based deep learning, generative models, computer animation, fluid simulation

## ACM Reference Format:

You Xie\*, Erik Franz\*, Mengyu Chu\*, and Nils Thuerey. 2018. tempoGAN: A Temporally Coherent, Volumetric GAN for Super-resolution Fluid Flow. *ACM Trans. Graph.* 37, 4, Article 95 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201304>

## 1 INTRODUCTION

Generative models were highly successful in the last years to represent and synthesize complex natural images [Goodfellow et al. 2014]. These works demonstrated that deep convolutional neural networks (CNNs) are able to capture the distribution of, e.g., photos of human faces, and generate novel, previously unseen versions that are virtually indistinguishable from the original inputs. Likewise, similar algorithms were shown to be extremely successful at generating natural high-resolution images from a coarse input [Karras et al. 2017]. However, in their original form, these generative models do not take into account the temporal evolution of the data, which is crucial for realistic physical systems. In the following, we will extend these methods to generate high-resolution volumetric data sets of passively advected flow quantities, and ensuring temporal coherence is one of the core aspects that we will focus on below. We

(\*) Similar amount of contributions.

Authors' addresses: You Xie\*, Technical University of Munich, [you.xie@tum.de](mailto:you.xie@tum.de); Erik Franz\*, Technical University of Munich, [franzer@in.tum.de](mailto:franzer@in.tum.de); Mengyu Chu\*, Technical University of Munich, [mengyu.chu@tum.de](mailto:mengyu.chu@tum.de); Nils Thuerey, Technical University of Munich, [nils.thuerey@tum.de](mailto:nils.thuerey@tum.de).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART95 \$15.00

<https://doi.org/10.1145/3197517.3201304>



# Fluid Directed Rigid Body Control using Deep Reinforcement Learning

PINGCHUAN MA, Nankai University, China

YUNSHENG TIAN\*, Nankai University, China

ZHERONG PAN, University of North Carolina at Chapel Hill, USA

BO REN†, Nankai University, China

DINESH MANOCHA, University of Maryland at College Park, USA

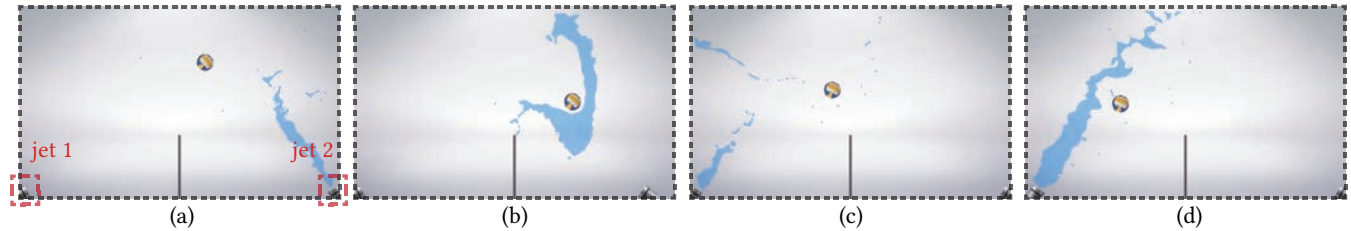


Fig. 1. Our new algorithm based on deep reinforcement learning can be used to control 2D coupled fluid/rigid animations. We use fluid jets (red blocks in (a)) to control a rigid body to exhibit desired behaviors. In this example, we use two fluid jets to play a cooperative ball game, where the goal is to push the ball back and forth. In these frames, the ball first flies to the right (a) and the right jet shoots the ball back (b). Next, the ball flies to the left (c), the left jet shoots it back (d). We use DRL to compute appropriate ghost forces at the fluid-solid boundary to control the ball's trajectory.

We present a learning-based method to control a coupled 2D system involving both fluid and rigid bodies. Our approach is used to modify the fluid/rigid simulator's behavior by applying control forces only at the simulation domain boundaries. The rest of the domain, corresponding to the interior, is governed by the Navier-Stokes equation for fluids and Newton-Euler's equation for the rigid bodies. We represent our controller using a general neural-net, which is trained using deep reinforcement learning. Our formulation decomposes a control task into two stages: a precomputation training stage and an online generation stage. We utilize various fluid properties, e.g., the liquid's velocity field or the smoke's density field, to enhance the controller's performance. We set up our evaluation benchmark by letting controller drive fluid jets move on the domain boundary and allowing them to shoot fluids towards a rigid body to accomplish a set of challenging 2D tasks such as keeping a rigid body balanced, playing a two-player ping-pong game, and driving a rigid body to sequentially hit specified points on the wall. In practice, our approach can generate physically plausible animations.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: fluid/rigid coupling, optimal control, reinforcement learning

\*joint first author

†corresponding author

Authors' addresses: Pingchuan Ma, Nankai University, 94 Weijin Rd, Tianjin, 300071, China, pika7ma@gmail.com; Yunsheng Tian, Nankai University, 94 Weijin Rd, Tianjin, 300071, China, knoxtian@gmail.com; Zherong Pan, University of North Carolina at Chapel Hill, Sitterson Hall, Chapel Hill, NC, 27514, USA, zherong@cs.unc.edu; Bo Ren, Nankai University, 94 Weijin Rd, Tianjin, 300071, China, rb@nankai.edu.cn; Dinesh Manocha, University of Maryland at College Park, A.V. Williams Building, 8223 Paint Branch Drive, College Park, MD, 20742, USA, dm@cs.unc.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART96 \$15.00

<https://doi.org/10.1145/3197517.3201334>

## ACM Reference Format:

Pingchuan Ma, Yunsheng Tian, Zherong Pan, Bo Ren, and Dinesh Manocha. 2018. Fluid Directed Rigid Body Control using Deep Reinforcement Learning. *ACM Trans. Graph.* 37, 4, Article 96 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201334>

## 1 INTRODUCTION

Fluid animation has been widely studied in computer graphics and related areas. Over the years, different techniques have been proposed to achieve visual plausibility on desktop machines. As a result, fluid simulators are used as standard tools in animation systems, e.g., Houdini and Naiad. Besides generating visually plausible animations, an additional requirement is for an artist to edit or control the fluid behaviors in a goal-directed or intuitive manner. Numerous methods, such as [Fattal and Lischinski 2004; McNamara et al. 2004; Pan et al. 2013; Pan and Manocha 2017; Shi and Yu 2005; Thuerney 2016; Treuille et al. 2003], have been proposed to this end.

The current control schemes do not account for the interactions between the fluid and other physical objects in the scene, such as solid boundaries or free-flying rigid objects. Ignoring these objects is reasonable for keyframe-based control applications where the goal is to deform the fluid into a target shape. However, in many other scenarios, the resulting simulation and control are based on fluid/rigid-body interactions. Typical applications of such scenarios include controlling a floating boat in a movie and controlling a volleyball using water jets in an AI game. In these scenarios, control is facilitated by applying forces on the fluid body, and the fluid body, in turn, influences the rigid body by applying forces and torques. Many prior algorithms [Fattal and Lischinski 2004; Shi and Yu 2005] apply virtual artificial control forces, or the so-called ghost forces, in the simulation domain to achieve controlling goals. Instead, we consider a much more challenging problem of boundary-condition-induced control, i.e. the controller only influences the fluid body at the simulation domain boundaries through physical actions

# StrokeAggregator: Consolidating Raw Sketches into Artist-Intended Curve Drawings

CHENXI LIU, University of British Columbia

ENRIQUE ROSALES, University of British Columbia and Universidad Panamericana

ALLA SHEFFER, University of British Columbia

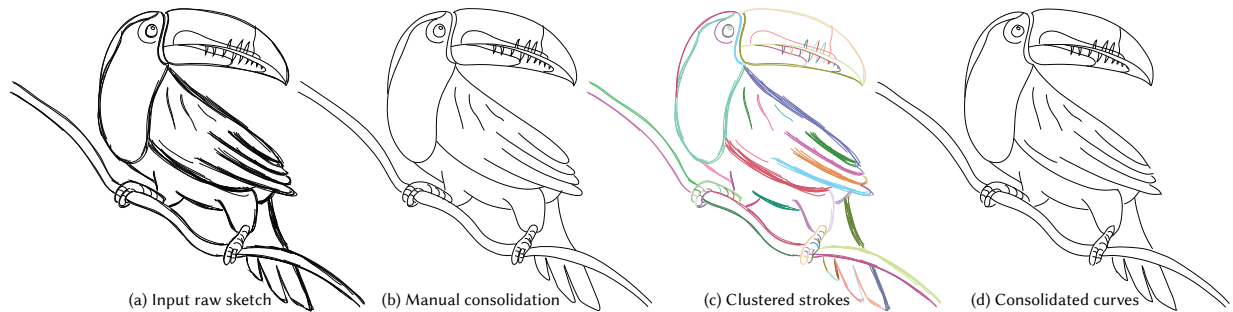


Fig. 1. Stroke consolidation: (a) a raw, vector format, sketch; (b) manually consolidated clean curve drawing; (c) algorithmically clustered strokes and (d) consolidated curves. Our output curve set (d) is of similar quality to the manually generated one (b). Please zoom in online to see image details. Raw sketch: © Enrique Rosales. Manual consolidation: © Elinor Palomares.

When creating line drawings, artists frequently depict intended curves using multiple, tightly clustered, or overdrawn, strokes. Given such sketches, human observers can readily envision these intended, *aggregate*, curves, and mentally assemble the artist's envisioned 2D imagery. Algorithmic stroke consolidation—replacement of overdrawn stroke clusters by corresponding aggregate curves—can benefit a range of sketch processing and sketch-based modeling applications which are designed to operate on consolidated, intended curves. We propose *StrokeAggregator*, a novel stroke consolidation method that significantly improves on the state of the art, and produces aggregate curve drawings validated to be consistent with viewer expectations. Our framework clusters strokes into groups that jointly define intended aggregate curves by leveraging principles derived from human perception research and observation of artistic practices. We employ these principles within a coarse-to-fine clustering method that starts with an initial clustering based on pairwise stroke compatibility analysis, and then refines it by analyzing interactions both within and in-between clusters of strokes. We facilitate this analysis by computing a common 1D parameterization for groups of strokes via common aggregate curve fitting. We demonstrate our method on a large range of line drawings, and validate its ability to generate consolidated drawings that are consistent with viewer perception via qualitative user evaluation, and comparisons to manually consolidated drawings and algorithmic alternatives.

Authors' addresses: Chenxi Liu, University of British Columbia, chenxil@cs.ubc.ca; Enrique Rosales, University of British Columbia, Universidad Panamericana, albertr@cs.ubc.ca; Alla Sheffer, University of British Columbia, sheffa@cs.ubc.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART97 \$15.00

<https://doi.org/10.1145/3197517.3201314>

CCS Concepts: • **Computing methodologies** → *Image manipulation*;

Additional Key Words and Phrases: Sketch consolidation, Gestalt perception, stroke clustering, curve fitting

## ACM Reference Format:

Chenxi Liu, Enrique Rosales, and Alla Sheffer. 2018. StrokeAggregator: Consolidating Raw Sketches into Artist-Intended Curve Drawings. *ACM Trans. Graph.* 37, 4, Article 97 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201314>

## 1 INTRODUCTION

Freehand line drawing provides a natural avenue for artists to quickly communicate shapes, ideas and images. When creating line drawings from scratch, artists often employ oversketching, using groups of multiple *raw* strokes to depict their intended, *aggregate*, curves (Figure 1a). Human observers can easily visually parse, or *consolidate*, these drawings by mentally replacing clusters of raw strokes with their corresponding aggregate curves. To create more refined, colorized, or shaded drawings, or to use these sketches as inputs to editing or modeling software, artists typically perform manual stroke consolidation by retracing the drawing and replacing raw stroke clusters with carefully drawn corresponding aggregate curves (Figure 1b) [Arora et al. 2017; Eissen and Steur 2008]. We present *StrokeAggregator*, a new algorithm that generates consolidated drawings of comparable quality to those generated by artists (Figure 1d).

Given the prevalence of tablets and other pen-sensitive displays, artists can easily create line drawings within a computer program and have the strokes recorded in vector form. These vector drawings contain more information about artist intent than their raster counterparts, motivating us to use vector format sketches as input. Algorithmic consolidation of both raster and vector drawings remains an open challenge: existing methods require parameter

# Real-Time Data-Driven Interactive Rough Sketch Inking

EDGAR SIMO-SERRA, Waseda University

SATOSHI IIZUKA, Waseda University

HIROSHI ISHIKAWA, Waseda University



Fig. 1. Example of interactive rough sketch inking. On the left we show an input rough sketch and on the right our user interface with the completed line drawing. The left of the user interface shows the canvas with user edits overlaid; green indicates *smart eraser* edits, light red shows *inker brush* edits, and dark red shows *inker pen* edits. On the right of the user interface we see the final result. This example was done using a mouse, and not any specialized illustration hardware. The image is copyrighted by Krenz Cushart and is part of the Krenz's Artwork Sketch Collection 2004-2013.

We present an interactive approach for *inking*, which is the process of turning a pencil rough sketch into a clean line drawing. The approach, which we call the **Smart Inker**, consists of several “smart” tools that intuitively react to user input, while guided by the input rough sketch, to efficiently and naturally connect lines, erase shading, and fine-tune the line drawing output. Our approach is data-driven: the tools are based on fully convolutional networks, which we train to exploit both the user edits and inaccurate rough sketch to produce accurate line drawings, allowing high-performance interactive editing in real-time on a variety of challenging rough sketch images. For the training of the tools, we developed two key techniques: one is the creation of training data by simulation of vague and quick user edits; the other is a line normalization based on learning from vector data. These techniques, in combination with our sketch-specific data augmentation, allow us to train the tools on heterogeneous data without actual user interaction. We validate our approach with an in-depth user study, comparing it with

professional illustration software, and show that our approach is able to reduce inking time by a factor of 1.8×, while improving the results of amateur users.

CCS Concepts: • **Computing methodologies** → **Image manipulation**; • **Human-centered computing** → *Gestural input*; • **Applied computing** → *Fine arts*;

Additional Key Words and Phrases: inking, interaction, sketching, line drawing

## ACM Reference Format:

Edgar Simo-Serra, Satoshi Iizuka, and Hiroshi Ishikawa. 2018. Real-Time Data-Driven Interactive Rough Sketch Inking. *ACM Trans. Graph.* 37, 4, Article 98 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201370>

## 1 INTRODUCTION

“1. The inker’s main purpose is to translate the penciller’s graphite pencil lines into reproducible, black, ink lines.

2. The inker must honor the penciller’s original intent while adjusting any obvious mistakes.

3. The inker determines the look of the finished art.”

— Gary Martin, *The Art of Comic Book Inking* [1997]

Although the role of computers in illustration has grown significantly in the last decade, pencil and paper sketching still remains the dominant way most illustrators brainstorm or begin new projects. Once a rough sketch is completed, it is usually inked, either digitally or by pen, to obtain a line drawing which is amenable to coloring, toning, or direct usage. The digital inking process consists of

Authors’ addresses: Edgar Simo-Serra, Waseda University, [esimo@aoni.waseda.jp](mailto:esimo@aoni.waseda.jp); Satoshi Iizuka, Waseda University, [iizuka@aoni.waseda.jp](mailto:iizuka@aoni.waseda.jp); Hiroshi Ishikawa, Waseda University, [hfs@waseda.jp](mailto:hfs@waseda.jp).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART98 \$15.00  
<https://doi.org/10.1145/3197517.3201370>



# FaceShop: Deep Sketch-based Face Image Editing

TIZIANO PORTENIER, University of Bern, Switzerland

QIYANG HU, University of Bern, Switzerland

ATTILA SZABÓ, University of Bern, Switzerland

SIAVASH ARJOMAND BIGDELI, University of Bern, Switzerland

PAOLO FAVARO, University of Bern, Switzerland

MATTHIAS ZWICKER, University of Maryland, College Park, USA



Fig. 1. Results produced by our sketch-based face image editing system. For each example, we show the original image and the input to our system in the left column, and the final output in the right column. Left: changing nose and strand of hair using sketching. Middle: editing eyes, makeup, and mouth using sketching and coloring. Right: replacing eyes region using smart copy-paste. Photos from the public domain.

We present a novel system for sketch-based face image editing, enabling users to edit images intuitively by sketching a few strokes on a region of interest. Our interface features tools to express a desired image manipulation by providing both geometry and color constraints as user-drawn strokes. As an alternative to the direct user input, our proposed system naturally supports a copy-paste mode, which allows users to edit a given image region by using parts of another exemplar image without the need of hand-drawn sketching at all. The proposed interface runs in real-time and facilitates an interactive and iterative workflow to quickly express the intended edits. Our system is based on a novel sketch domain and a convolutional neural network trained end-to-end to automatically learn to render image regions corresponding to the input strokes. To achieve high quality and semantically consistent results we train our neural network on two simultaneous tasks, namely image completion and image translation. To the best of our knowledge, we are the first to combine these two tasks in a unified framework for interactive image editing. Our results show that the proposed sketch domain, network architecture, and training procedure generalize well to real user input and enable high quality synthesis results without additional post-processing.

Authors' addresses: Tiziano Portenier, University of Bern, Neubrückstrasse 10, Bern, 3012, Switzerland, [portenier@inf.unibe.ch](mailto:portenier@inf.unibe.ch); Qiyang Hu, University of Bern, Neubrückstrasse 10, Bern, 3012, Switzerland, [hu@inf.unibe.ch](mailto:hu@inf.unibe.ch); Attila Szabó, University of Bern, Neubrückstrasse 10, Bern, 3012, Switzerland, [szabo@inf.unibe.ch](mailto:szabo@inf.unibe.ch); Siavash Arjomand Bigdeli, University of Bern, Neubrückstrasse 10, Bern, 3012, Switzerland, [bigdeli@inf.unibe.ch](mailto:bigdeli@inf.unibe.ch); Paolo Favaro, University of Bern, Neubrückstrasse 10, Bern, 3012, Switzerland, [favaro@inf.unibe.ch](mailto:favaro@inf.unibe.ch); Matthias Zwicker, University of Maryland, College Park, 8223 Paint Branch Drive, College Park, MD, 20742, USA, [zwicker@cs.umd.edu](mailto:zwicker@cs.umd.edu).



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART99  
<https://doi.org/10.1145/3197517.3201393>

CCS Concepts: • **Computing methodologies** → **Image manipulation**;

Additional Key Words and Phrases: Sketch-based interface, convolutional neural network, image editing

## ACM Reference Format:

Tiziano Portenier, Qiyang Hu, Attila Szabó, Siavash Arjomand Bigdeli, Paolo Favaro, and Matthias Zwicker. 2018. FaceShop: Deep Sketch-based Face Image Editing. *ACM Trans. Graph.* 37, 4, Article 99 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201393>

## 1 INTRODUCTION

Interactive image editing is an important field in both computer graphics and computer vision. Social media platforms register rapidly growing contributions of imagery content, which is creating an increasing demand for image editing applications that enable untrained users to enhance and manipulate photos. For example, Instagram reports more than 40 billion uploaded photos, and the number of photos is growing by 95 million per day [Instagram 2018]. However, there is a lack of tools that feature more complex editing operations for inexperienced users, such as changing the facial expression in an image. In this work, we propose a sketch-based editing framework that enables a user to edit a given image by sketching a few strokes on top of it. We show three examples produced by our system in Figure 1 and more results in Section 4.

Image editing techniques can be divided into global and local editing operations. Recently, numerous deep learning techniques have been proposed for global editing applications such as general image enhancing [Gharbi et al. 2017; Yan et al. 2016], grayscale to color transformation [Iizuka et al. 2016; Zhang et al. 2017], and geometric, illumination, and style transformations [Gatys et al. 2015;

# Space-time Tomography for Continuously Deforming Objects

GUANGMING ZANG, RAMZI IDOUGHI, RAN TAO, GILLES LUBINEAU, PETER WONKA, and WOLFGANG HEIDRICH, King Abdullah University of Science And Technology, Saudi Arabia

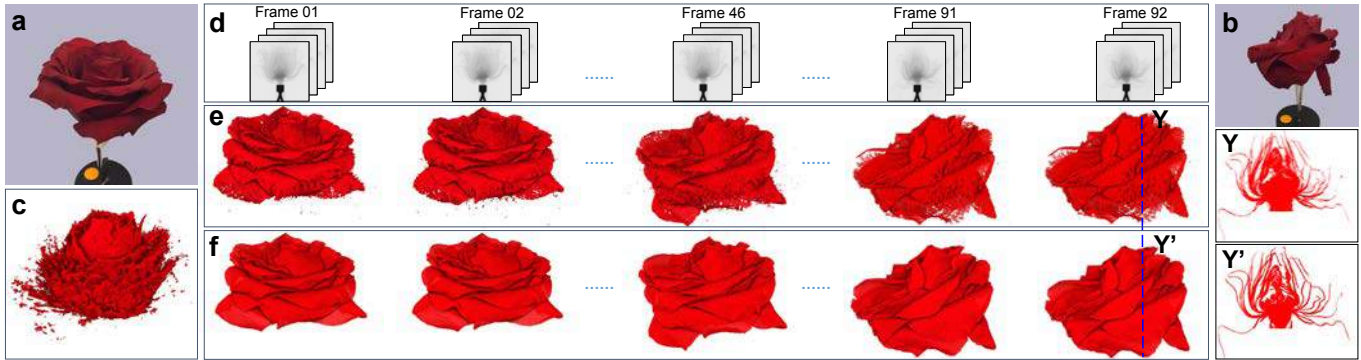


Fig. 1. X-ray tomographic reconstruction of a rose undergoing significant wilting during the scan due to loss of water. Images (a) and (b) show photographs of the rose directly before and directly after the scanning process. Traditional CT reconstruction (c) from all 5520 projections in the scan sequence show significant distortions due to misalignment of features. When grouping the projections into 92 frames of 60 projections each (d), the deformation over each frame becomes negligible, but now the number of projections per frame is insufficient for high-quality reconstruction of the corresponding volumes (e, Y). By comparison, our full space-time reconstruction algorithm yields a time sequence of highly detailed volumes for different time steps (f, Y').

X-ray computed tomography (CT) is a valuable tool for analyzing objects with interesting internal structure or complex geometries that are not accessible with optical means. Unfortunately, tomographic reconstruction of complex shapes requires a multitude (often hundreds or thousands) of projections from different viewpoints. Such a large number of projections can only be acquired in a time-sequential fashion. This significantly limits the ability to use x-ray tomography for either objects that undergo uncontrolled shape change at the time scale of a scan, or else for analyzing dynamic phenomena, where the motion itself is under investigation.

In this work, we present a non-parametric space-time tomographic method for tackling such dynamic settings. Through a combination of a new CT image acquisition strategy, a space-time tomographic image formation model, and an alternating, multi-scale solver, we achieve a general approach that can be used to analyze a wide range of dynamic phenomena. We demonstrate our method with extensive experiments on both real and simulated data.

CCS Concepts: • **Computing methodologies** → *Computer graphics; 3D imaging; Motion capture*;

Additional Key Words and Phrases: X-ray computed tomography, 4D reconstruction, Optimization

Authors' address: Guangming Zang, [guangming.zang@kaust.edu.sa](mailto:guangming.zang@kaust.edu.sa); Ramzi Idoughi, [ramzi.idoughi@kaust.edu.sa](mailto:ramzi.idoughi@kaust.edu.sa); Ran Tao, [ran.tao@kaust.edu.sa](mailto:ran.tao@kaust.edu.sa); Gilles Lubineau, [gilles.lubineau@kaust.edu.sa](mailto:gilles.lubineau@kaust.edu.sa); Peter Wonka, [peter.wonka@kaust.edu.sa](mailto:peter.wonka@kaust.edu.sa); Wolfgang Heidrich, [wolfgang.heidrich@kaust.edu.sa](mailto:wolfgang.heidrich@kaust.edu.sa), King Abdullah University of Science And Technology, Thuwal, 23955-6900, Saudi Arabia.



This work is licensed under a Creative Commons Attribution-NonCommercial International 4.0 License.

© 2018 Copyright held by the owner/author(s).  
0730-0301/2018/8-ART100  
<https://doi.org/10.1145/3197517.3201298>

## ACM Reference Format:

Guangming Zang, Ramzi Idoughi, Ran Tao, Gilles Lubineau, Peter Wonka, and Wolfgang Heidrich. 2018. Space-time Tomography for Continuously Deforming Objects. *ACM Trans. Graph.* 37, 4, Article 100 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201298>

## 1 INTRODUCTION

Capture of 3D geometry and dynamic phenomena has long been a topic of interest in computer graphics. While much of the work in this area has been based on traditional computer vision methods with either passive or active illumination [Dou et al. 2016; Józsa et al. 2013], tomographic methods have also played an important role in computer graphics research [Gregson et al. 2014; Ijiri et al. 2014; Reche-Martinez et al. 2004; Stuppy et al. 2003; Zhao et al. 2011].

In engineering and medicine, x-ray computed tomography has become a valuable tool for its ability to reveal internal structures in objects that cannot be sensed by optical means. Tomographic reconstruction in graphics has been primarily focused on visible light tomography [Reche-Martinez et al. 2004; Trifonov et al. 2006; Zheng et al. 2017], although there are a few examples of using x-ray CTs as well [Stuppy et al. 2003; Zhao et al. 2011]. While visible light tomography is inherently limited to transparent materials, it does have the advantage of using regular cameras and light sources, which makes it feasible to construct multi-camera setups that can be used to capture dynamic phenomena such as fluids (e.g. [Atcheson et al. 2008; Elsinga et al. 2006; Gregson et al. 2014, 2012; Hasinoff and Kutulakos 2007; Ihrke and Magnor 2004]). By comparison, x-ray CT is typically limited to a single source and sensor array or a very small number of such sources and arrays, and the required diversity of view points is created through mechanical motion of either the sample (typical of industrial CTs) or the source-sensor assembly (as

# Instant 3D Photography

PETER HEDMAN, University College London\*

JOHANNES KOPF, Facebook

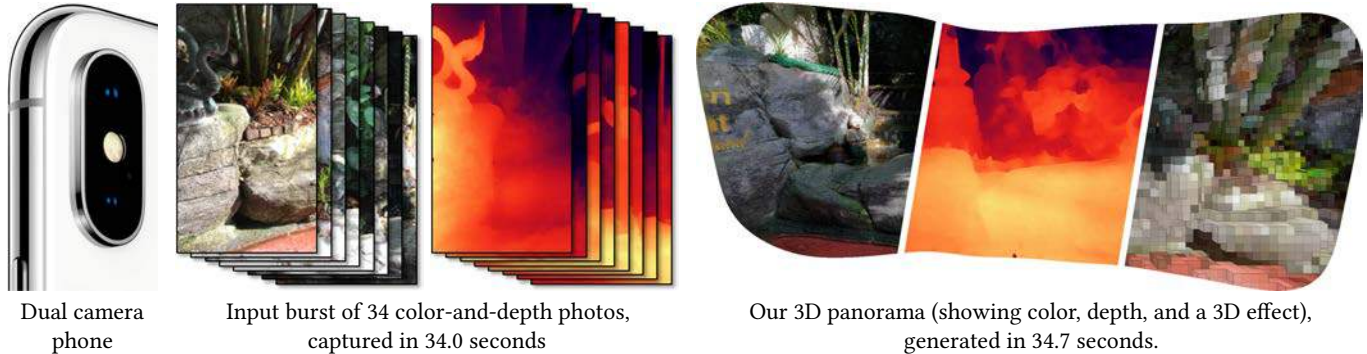


Fig. 1. Our work enables practical and casual 3D capture with regular dual camera cell phones. *Left*: A burst of input color-and-depth image pairs that we captured with a dual camera cell phone at a rate of one image per second. *Right*: 3D panorama generated with our algorithm in about the same time it took to capture. The geometry is highly detailed and enables viewing with binocular and motion parallax in VR, as well as applying 3D effects that interact with the scene, e.g., through occlusions (right).

We present an algorithm for constructing 3D panoramas from a sequence of aligned color-and-depth image pairs. Such sequences can be conveniently captured using dual lens cell phone cameras that reconstruct depth maps from synchronized stereo image capture. Due to the small baseline and resulting triangulation error the depth maps are considerably degraded and contain low-frequency error, which prevents alignment using simple global transformations. We propose a novel optimization that jointly estimates the camera poses as well as spatially-varying adjustment maps that are applied to deform the depth maps and bring them into good alignment. When fusing the aligned images into a seamless mosaic we utilize a carefully designed data term and the high quality of our depth alignment to achieve two orders of magnitude speedup w.r.t. previous solutions that rely on discrete optimization by removing the need for label smoothness optimization. Our algorithm processes about one input image per second, resulting in an end-to-end runtime of about one minute for mid-sized panoramas. The final 3D panoramas are highly detailed and can be viewed with binocular and head motion parallax in VR.

CCS Concepts: • **Computing methodologies** → **Image-based rendering**; **Reconstruction**; **Computational photography**;

Additional Key Words and Phrases: 3D Reconstruction, Image-stitching, Image-based Rendering, Virtual Reality

\* This work was done while Peter was working as a contractor for Facebook.

Authors' addresses: Peter Hedman, University College London\*, p.hedman@cs.ucl.ac.uk; Johannes Kopf, Facebook, jkopf@fb.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART101 \$15.00

<https://doi.org/10.1145/3197517.3201384>

## 1 INTRODUCTION

Virtual reality (VR) is a fascinating emerging technology that creates lifelike experiences in immersive virtual environments, with high-end headsets now widely available. Most content that is consumed in VR today is synthetic and needs to be created by professional artists. There is no practical way for consumers to capture and share their own real-life environments in a form that makes full use of the VR technology.

That is perhaps not surprising when considering the formidability of this problem: we are looking for a method that does not require expensive hardware and is easy to use even for novice users. Yet, it should create high-quality and *truly immersive* content, i.e., a 3D representation that supports binocular vision and head-motion parallax. Finally, consumers demand very fast processing times, on the order of seconds at most.

Panoramic images and video can be easily captured now with consumer 360° cameras. While the surrounding imagery provides some immersion, the realism is limited due to lack of depth and parallax. Stereo panoramas [Peleg and Ben-Ezra 1999] provide binocular depth cues by delivering different images to the left and right eye, but these images are static and do not provide motion parallax when the user turns or moves their head.

Full immersion can only be achieved using a 3D representation, such as the 3D models generated by multi-view stereo methods. However, when applied to typical panorama datasets, captured “inside-out” from a single vantage point, these methods have to deal with a very small baseline, which causes noisy results and holes in the reconstruction. The methods are also very sensitive to even slight scene motion, such as wind-induced motion in trees. The *Casual 3D Photography* system [Hedman et al. 2017] achieves some improvement in reconstruction quality, but it is slow with a runtime of several hours per scene.



# Reconstructing Scenes with Mirror and Glass Surfaces

THOMAS WHELAN, Facebook Reality Labs

MICHAEL GOESELE, Facebook Reality Labs and TU Darmstadt\*

STEVEN J. LOVEGROVE, JULIAN STRAUB, and SIMON GREEN, Facebook Reality Labs

RICHARD SZELISKI, Facebook

STEVEN BUTTERFIELD, SHOBHIT VERMA, and RICHARD NEWCOMBE, Facebook Reality Labs

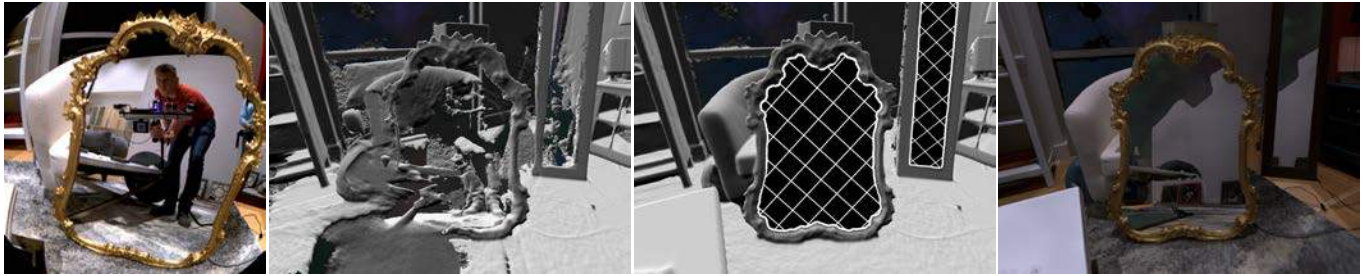


Fig. 1. Reconstructing a scene with mirrors. *From left to right:* Input color image showing the scanner with attached AprilTag in a mirror, reconstructed geometry without taking the mirrors into account, reconstruction taking the detected mirrors (rendered as cross-hatched area) into account and a photorealistic rendering of the scene including the mirrors. Detecting the mirrors is crucial for accurate geometry reconstruction and realistic rendering.

Planar reflective surfaces such as glass and mirrors are notoriously hard to reconstruct for most current 3D scanning techniques. When treated naïvely, they introduce duplicate scene structures, effectively destroying the reconstruction altogether. Our key insight is that an easy to identify structure attached to the scanner—in our case an AprilTag—can yield reliable information about the existence and the geometry of glass and mirror surfaces in a scene. We introduce a fully automatic pipeline that allows us to reconstruct the geometry and extent of planar glass and mirror surfaces while being able to distinguish between the two. Furthermore, our system can automatically segment observations of multiple reflective surfaces in a scene based on their estimated planes and locations. In the proposed setup, minimal additional hardware is needed to create high-quality results. We demonstrate this using reconstructions of several scenes with a variety of real mirrors and glass.

CCS Concepts: • **Computing methodologies** → **3D imaging; Shape modeling; Reflectance modeling;**

Additional Key Words and Phrases: 3D scanning, reflective surfaces, mirrors, glass

\*This work was carried out at Facebook Reality Labs.

Authors' addresses: Thomas Whelan, Facebook Reality Labs, [twhelan@fb.com](mailto:twhelan@fb.com); Michael Goesele, Facebook Reality Labs and TU Darmstadt; Steven J. Lovegrove; Julian Straub; Simon Green, Facebook Reality Labs; Richard Szeliski, Facebook; Steven Butterfield; Shobhit Verma; Richard Newcombe, Facebook Reality Labs.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART102 \$15.00  
<https://doi.org/10.1145/3197517.3201319>

## ACM Reference Format:

Thomas Whelan, Michael Goesele, Steven J. Lovegrove, Julian Straub, Simon Green, Richard Szeliski, Steven Butterfield, Shobhit Verma, and Richard Newcombe. 2018. Reconstructing Scenes with Mirror and Glass Surfaces. *ACM Trans. Graph.* 37, 4, Article 102 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201319>

## 1 INTRODUCTION

Active scanning techniques using Kinect-like sensors have recently been used very successfully to reconstruct indoor scenes [Kähler et al. 2015; Newcombe et al. 2011; Nießner et al. 2013]. One frequently occurring element of such scenes—reflective surfaces such as mirrors and glass—has, rarely been treated in previous work. While this at first sounds like a minor omission, mirrors actually pose a significant problem for any reconstruction system. A perfect mirror shows a perfect reflection of the world, which is indistinguishable from a direct observation of the mirrored world. The mirror is therefore essentially “invisible”. The mirrored scene, however, will still be reconstructed using standard vision techniques. Fig. 2 shows examples of bathrooms from the ScanNet dataset [Dai et al. 2017] with reflected scene parts reconstructed behind mirrors.

This mirrored geometry overlaps with real geometry located behind the mirror and interferes with its reconstruction. It is therefore not only desirable to reconstruct mirrors as standard elements in a scene, but also necessary in order to avoid invalid reconstructions of the overall scene. Similarly, glass surfaces are typically not reconstructed by sensors but should still be included in a reconstructed model. Due to the prevalence of mirrors and glass in common indoor environments, recent scene reconstruction approaches such as Matterport3D [Chang et al. 2017] require the user to manually select windows and mirrors in a scan.

# Full 3D Reconstruction of Transparent Objects

BOJIAN WU, SIAT, Shenzhen University, and University of Chinese Academy of Sciences

YANG ZHOU, Shenzhen University and Huazhong University of Science & Technology

YIMING QIAN, University of Alberta

MINGLUN GONG, Memorial University of Newfoundland

HUI HUANG\*, Shenzhen University

Numerous techniques have been proposed for reconstructing 3D models for opaque objects in past decades. However, none of them can be directly applied to transparent objects. This paper presents a fully automatic approach for reconstructing complete 3D shapes of transparent objects. Through positioning an object on a turntable, its silhouettes and light refraction paths under different viewing directions are captured. Then, starting from an initial rough model generated from space carving, our algorithm progressively optimizes the model under three constraints: surface and refraction normal consistency, surface projection and silhouette consistency, and surface smoothness. Experimental results on both synthetic and real objects demonstrate that our method can successfully recover the complex shapes of transparent objects and faithfully reproduce their light refraction properties.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Shape modeling**; **Point-based models**;

Additional Key Words and Phrases: 3D reconstruction, transparent Objects

## ACM Reference Format:

Bojian Wu, Yang Zhou, Yiming Qian, Minglun Gong, and Hui Huang. 2018. Full 3D Reconstruction of Transparent Objects. *ACM Trans. Graph.* 37, 4, Article 103 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201286>

## 1 INTRODUCTION

Reconstructing 3D models of real objects has been an active research topic in both Computer Vision and Graphics for decades. A variety of approaches have been proposed for different applications, such as autonomous scanning [Wu et al. 2014], multi-view stereo [Galliani et al. 2015], photometric stereo [Chen et al. 2007], etc. While these techniques are able to faithfully capture and reconstruct the shapes of opaque or even translucent objects, none of them can be directly applied on transparent objects. As a result, people often have to paint those transparent objects before capturing their shapes.

On another front, how transparent objects refract lights toward a fixed viewpoint can be accurately acquired using environment matting techniques [Chuang et al. 2000; Qian et al. 2015]. Since



Fig. 1. A transparent object refracts lights in its environment (left) and hence its shape cannot be reconstructed using conventional techniques. We present a novel method that, for the first time, directly reconstruct the shapes of transparent objects from captured images. This allows photo-realistic rendering of real transparent objects in virtual environments (right). For a clear view of the reconstructed geometry, please see Fig. 17.

light refraction paths are determined by surface normals, one has to wonder whether the shape of the transparent object can be inferred accordingly. This has been demonstrated as a feasible direction in a previous work [Qian et al. 2016]. Through enforcing a position-normal consistency constraint, their approach can generate point clouds on two sides of a given transparent object. Nevertheless, the captured surface shape is incomplete. In addition, this approach is not that easy to apply, involving setting up and calibrating among 4 different camera/monitor configurations.

This paper presents a fully automatic approach for reconstructing complete 3D shapes of transparent objects with known refractive indexes. Through positioning the object on a turntable, its silhouettes and light refraction paths under different viewing directions are captured using two fixed cameras (Section 3). An initial model generated through space carving is then gradually evolved toward the accurate object shape using novel point consolidation formulations that are constrained by captured light refraction paths and silhouettes (Section 4). Results on both synthetic and real objects (Section 5) demonstrate the effectiveness and robustness of our approach; see e.g., Fig. 1.

## 2 RELATED WORK

**Surface reconstruction.** The literature on 3D surface reconstruction [Berger et al. 2013, 2014] is vast. Specifically, reconstructing transparent objects is well-known a challenging problem [Ihrke et al. 2010]. Recent developments, such as reconstruction of flames [Ihrke and Magnor 2004; Wu et al. 2015b], mixing fluids [Gregson et al.

\*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Bojian Wu, SIAT, Shenzhen University, University of Chinese Academy of Sciences; Yang Zhou, Shenzhen University, Huazhong University of Science & Technology; Yiming Qian, University of Alberta; Minglun Gong, Memorial University of Newfoundland; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART103 \$15.00

<https://doi.org/10.1145/3197517.3201286>

# Object-Aware Guidance for Autonomous Scene Reconstruction

LIGANG LIU, University of Science and Technology of China

XI XIA, University of Science and Technology of China

HAN SUN, University of Science and Technology of China

QI SHEN, University of Science and Technology of China

JUZHAN XU, Shenzhen University

BIN CHEN, Shenzhen University

HUI HUANG, Shenzhen University

KAI XU\*, Shenzhen University and National University of Defense Technology

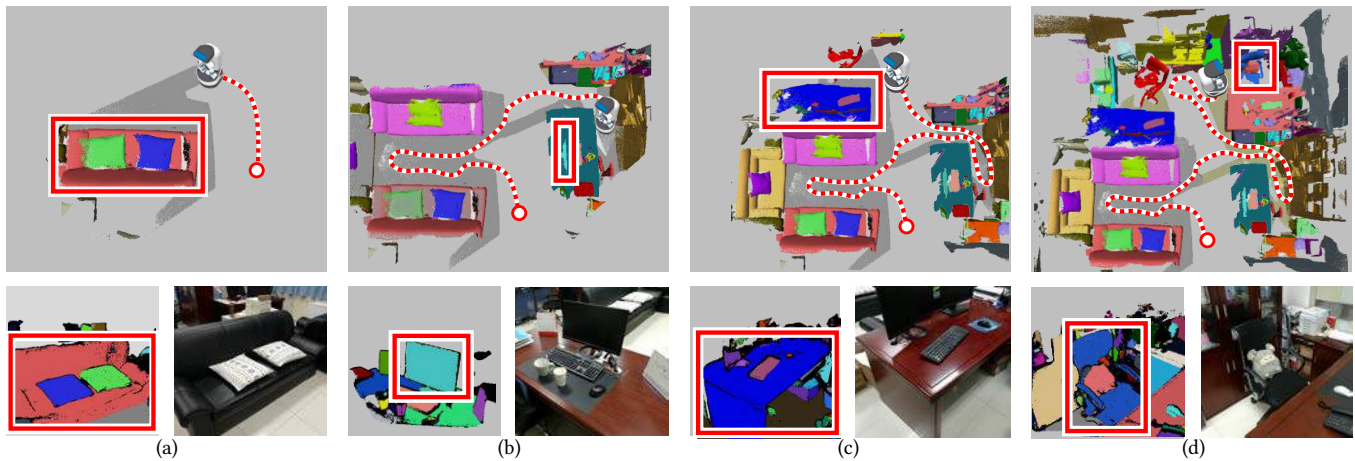


Fig. 1. Autonomous scene scanning and reconstruction on a real office scene using our object-aware guidance approach. In each column (a)-(d), the object marked with the red rectangular frame is the object-of-interest (OOI). The upper row shows the navigation path (in dotted red) with previous scanning views (shown as white dots) and the current position of the robot. The objects in different colors are the reconstructed objects in the scene. The bottom row shows the depth data (left) and the RGB image (right) from the current view of the robot. Our approach achieves both global path planning and local view planning on-the-fly within one single navigation pass and obtains the reconstructed scene with semantic objects (d).

To carry out autonomous 3D scanning and online reconstruction of unknown indoor scenes, one has to find a balance between global exploration of the entire scene and local scanning of the objects within it. In this work, we propose a novel approach, which provides object-aware guidance for autoscanning, for exploring, reconstructing, and understanding an unknown scene within *one navigation pass*. Our approach interleaves between object analysis to identify the *next best object* (NBO) for global exploration, and object-aware information gain analysis to plan the *next best view* (NBV) for

local scanning. First, an objectness-based segmentation method is introduced to extract semantic objects from the current scene surface via a multi-class graph cuts minimization. Then, an object of interest (OOI) is identified as the NBO which the robot aims to visit and scan. The robot then conducts fine scanning on the OOI with views determined by the NBV strategy. When the OOI is recognized as a full object, it can be replaced by its most similar 3D model in a shape database. The algorithm iterates until all of the objects are recognized and reconstructed in the scene. Various experiments and comparisons have shown the feasibility of our proposed approach.

CCS Concepts: • **Computing methodologies** → *Shape analysis*;

Additional Key Words and Phrases: Indoor scene reconstruction, autonomous reconstruction, next-best-object, next-best-view

## ACM Reference Format:

Ligang Liu, Xi Xia, Han Sun, Qi Shen, Juzhan Xu, Bin Chen, Hui Huang, and Kai Xu. 2018. Object-Aware Guidance for Autonomous Scene Reconstruction. *ACM Trans. Graph.* 37, 4, Article 104 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201295>

## 1 INTRODUCTION

In recent years autonomous 3D scanning and reconstruction of indoor scenes by mobile robots with RGB-D depth sensors have

\*Kai Xu is the corresponding author (kevin.kai.xu@gmail.com)

Authors' addresses: Ligang Liu, University of Science and Technology of China; Xi Xia, University of Science and Technology of China; Han Sun, University of Science and Technology of China; Qi Shen, University of Science and Technology of China; Juzhan Xu, Shenzhen University; Bin Chen, Shenzhen University; Hui Huang, Shenzhen University; Kai Xu, Shenzhen University, National University of Defense Technology.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART104 \$15.00

<https://doi.org/10.1145/3197517.3201295>



# Optimal Cone Singularities for Conformal Flattening

YOUSUF SOLIMAN, DEJAN SLEPČEV, and KEENAN CRANE, Carnegie Mellon University

Angle-preserving or *conformal* surface parameterization has proven to be a powerful tool across applications ranging from geometry processing, to digital manufacturing, to machine learning, yet conformal maps can still suffer from severe area distortion. *Cone singularities* provide a way to mitigate this distortion, but finding the best configuration of cones is notoriously difficult. This paper develops a strategy that is globally optimal in the sense that it minimizes total area distortion among all possible cone configurations (number, placement, and size) that have no more than a fixed total cone angle. A key insight is that, for the purpose of optimization, one should not work directly with curvature measures (which naturally represent cone configurations), but can instead apply *Fenchel-Rockafellar duality* to obtain a formulation involving only ordinary functions. The result is a convex optimization problem, which can be solved via a sequence of sparse linear systems easily built from the usual cotangent Laplacian. The method supports user-defined notions of importance, constraints on cone angles (e.g., positive, or within a given range), and sophisticated boundary conditions (e.g., convex, or polygonal). We compare our approach to previous techniques on a variety of challenging models, often achieving dramatically lower distortion, and demonstrating that global optimality leads to extreme robustness in the presence of noise or poor discretization.

CCS Concepts: • **Computing methodologies** → **Mesh geometry models**; • **Mathematics of computing** → *Continuous optimization*;

Additional Key Words and Phrases: discrete differential geometry, geometry processing, conformal mapping, mesh parameterization

## ACM Reference Format:

Yousuf Soliman, Dejan Slepčev, and Keenan Crane. 2018. Optimal Cone Singularities for Conformal Flattening. *ACM Trans. Graph.* 37, 4, Article 105 (August 2018), 24 pages. <https://doi.org/10.1145/3197517.3201367>

## 1 INTRODUCTION

Mesh parameterization is a fundamental component of a wide variety of problems in applied geometry: beyond traditional tasks in computer graphics (such as texture mapping), surface flattenings have become an important component in a diverse collection of areas ranging from digital manufacturing to machine learning [Konakovic et al. 2016; Maron et al. 2017]. Ideally, one would like a parameterization that is *isometric*, i.e., no distortion of lengths or areas, but for general curved surfaces no such map exists. Conformal flattenings are attractive because they completely eliminate angle distortion, and are easily computed via linear or convex problems. However, they can also yield significant distortion of areas, which is problematic for applications since a large region of the surface

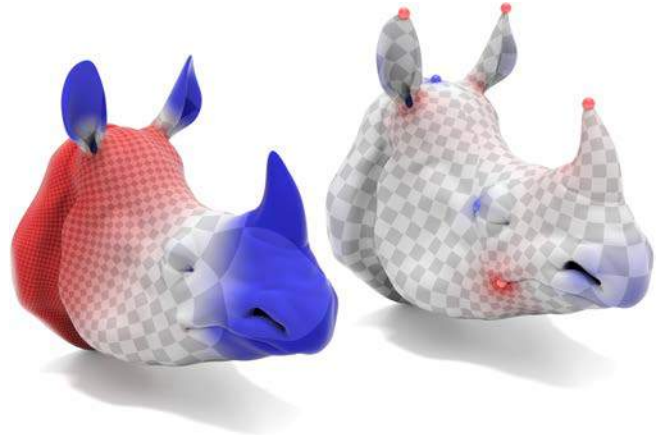


Fig. 1. *Left*: Conformal flattenings have no angle distortion, but can severely distort area. *Right*: When placed optimally, even a small number of cone singularities (here just nine) can almost completely eliminate area distortion.

is represented by only a tiny region in the parameter domain. The basic idea behind *cone flattening* [Kharevych et al. 2006] is that, intrinsically, many surfaces look more like a polyhedron than the flat plane—consider for instance maps of the Earth generated by conformally mapping the globe onto a regular polyhedron (Fig. 2). Since this initial map induces very little area distortion, and since the polyhedron can then be cut and unfolded into the plane without further stretching, the composite map also has low area distortion. Of course, different polyhedral metrics will lead to different amounts of distortion—the problem of cone parameterization therefore boils down to deciding on a configuration of vertices, determined by the number, placement, and angle of the associated *cone singularities*. As stated, however, this problem is ill-posed: one can always reduce distortion further by considering a finer polyhedron, i.e., allowing more cones. To make the problem well-posed, one can fix the number of cones, or alternatively (as we will do), fix the total magnitude  $\Phi := \sum_i |\phi_i|$  of all cone angles  $\phi_i$ .



Fig. 2. A conformal cone parameterization is equivalent to flattening a smooth surface, like the sphere, over a polyhedron, which can then be cut and unfolded into the plane without further distortion. By adding more and more cone points, one can make area distortion arbitrarily small. (Texture courtesy NASA Earth Observatory.)

Authors' address: Yousuf Soliman; Dejan Slepčev; Keenan Crane, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART105 \$15.00 <https://doi.org/10.1145/3197517.3201367>

# Rapid Deployment of Curved Surfaces via Programmable Auxetics

MINA KONAKOVIĆ-LUKOVIĆ, EPFL

JULIAN PANETTA, EPFL

KEENAN CRANE, Carnegie Mellon University

MARK PAULY, EPFL

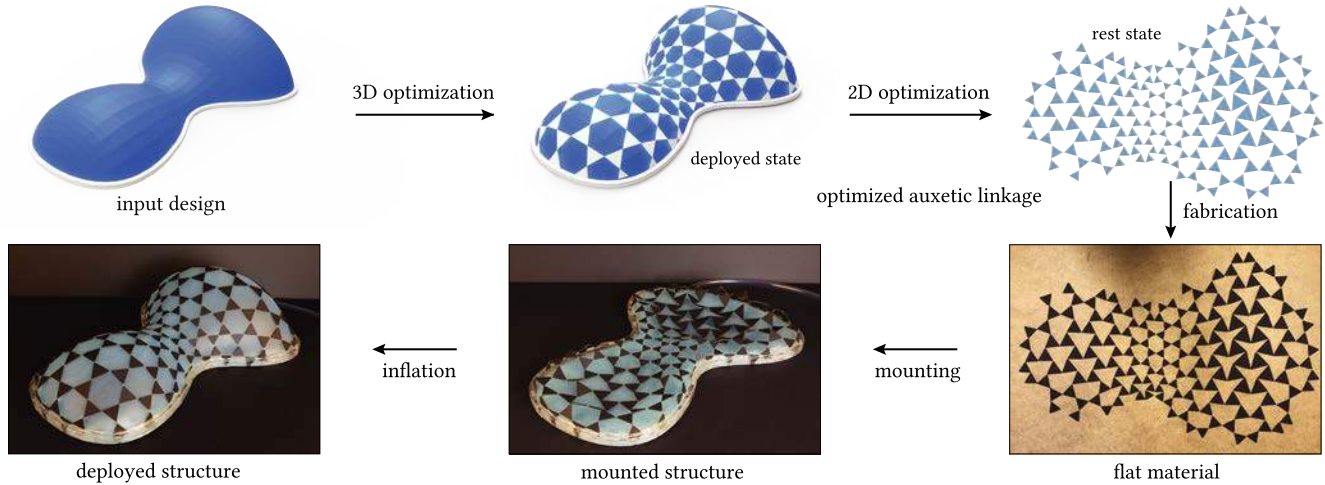


Fig. 1. Our algorithm computes a triangular auxetic linkage that closely approximates a given surface when deployed to maximal extension via inflation. The fabricated material is laser cut from a single sheet, mounted onto the support frame, and inflated with a generic rubber balloon.

Deployable structures are physical mechanisms that can easily transition between two or more geometric configurations; such structures enable industrial, scientific, and consumer applications at a wide variety of scales. This paper develops novel deployable structures that can approximate a large class of doubly-curved surfaces and are easily actuated from a flat initial state via inflation or gravitational loading. The structures are based on two-dimensional rigid mechanical linkages that implicitly encode the curvature of the target shape via a user-programmable pattern that permits locally isotropic scaling under load. We explicitly characterize the shapes that can be realized by such structures—in particular, we show that they can approximate target surfaces of positive mean curvature and bounded scale distortion relative to a given reference domain. Based on this observation, we develop efficient computational design algorithms for approximating a given input geometry. The resulting designs can be rapidly manufactured via digital fabrication technologies such as laser cutting, CNC milling, or 3D printing. We validate our approach through a series of physical prototypes and present several application case studies, ranging from surgical implants to large-scale deployable architecture.

Authors' addresses: Mina Konaković-Luković, mina.konakovic@epfl.ch, EPFL, Route Cantonale, Lausanne, Switzerland, 1015; Julian Panetta, EPFL, julian.panetta@epfl.ch; Keenan Crane, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, Pennsylvania, 15213, kncrane@cs.cmu.edu; Mark Pauly, EPFL, mark.pauly@epfl.ch.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART106 \$15.00

<https://doi.org/10.1145/3197517.3201373>

CCS Concepts: • **Computing methodologies** → *Mesh geometry models*; • **Applied computing** → *Computer-aided design*; *Computer-aided manufacturing*;

Additional Key Words and Phrases: computational fabrication, smart materials, digital fabrication, auxetic materials, conformal geometry

## ACM Reference Format:

Mina Konaković-Luković, Julian Panetta, Keenan Crane, and Mark Pauly. 2018. Rapid Deployment of Curved Surfaces via Programmable Auxetics. *ACM Trans. Graph.* 37, 4, Article 106 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201373>

## 1 INTRODUCTION

Deployable structures are shape-shifting mechanisms that can transition between two or more geometric configurations. Often conceived to minimize space requirements for storage or transport, deployable structures are used, for example, for antennas or solar panels in satellites, as coronary stents in medical applications, as consumer products (e.g. umbrellas), or in architectural designs (e.g. retractable bridges or relocatable, temporary event spaces).

Most existing realizations of deployable structures are geometrically simple and often exhibit strong symmetries. Deploying more general curved surfaces is made difficult by the inherent complexity of jointly designing initial and target geometries within the constraints imposed by the deployment mechanism [Gantes 2001].

We propose a new class of deployable structures and associated computational methods that enable rapid deployment of doubly-curved freeform surfaces (see Figure 1). Our approach is based on a

# Designing Patterns using Triangle-Quad Hybrid Meshes

CHI-HAN PENG, KAUST  
HELMUT POTTMANN, TU Wien  
PETER WONKA, KAUST

We present a framework to generate mesh patterns that consist of a *hybrid* of both triangles and quads. Given a 3D surface, the generated patterns fit the surface boundaries and curvatures. Such regular and near regular triangle-quad hybrid meshes provide two key advantages: first, novel-looking polygonal patterns achieved by mixing different arrangements of triangles and quads together; second, a finer discretization of angle deficits than utilizing triangles or quads alone. Users have controls over the generated patterns in global and local levels. We demonstrate applications of our approach in architectural geometry and pattern design on surfaces.

CCS Concepts: • **Computing methodologies** → **Mesh models; Mesh geometry models**;

Additional Key Words and Phrases: Meshes, Polygonal Patterns, Pattern Design, Geometric Modeling

## ACM Reference Format:

Chi-Han Peng, Helmut Pottmann, and Peter Wonka. 2018. Designing Patterns using Triangle-Quad Hybrid Meshes. *ACM Trans. Graph.* 37, 4, Article 107 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201306>

## 1 INTRODUCTION

The study of tilings and patterns has various applications, e.g. in architecture, industrial design, and art. In this work we study patterns that are generated from triangle- and quad-shaped tiles. One motivation for our work is that there is a large body of research studying the meshing of surfaces using triangle, quad, or hex meshes. In addition, there are also quad-dominant meshes, that mainly consist of quads with a few isolated triangles. In this paper, we would like to explore a richer class of designs that are not confined to be either triangle or quad(-dominant) meshes, but a *hybrid* of both. See Fig. 1 for a novel architectural mesh design enabled by our work.

To start our investigation, we first look at meshes in the plane consisting of regular triangles and squares. In this context, the tiling of the infinite Euclidean plane has been studied for various tile sets, including tile sets using quads and triangles [Grunbaum and Shephard 2016]). However, an open question is how to tile a finite domain with a given boundary using squares and regular triangles. As an important building block of our work we analyze this tiling problem and propose an optimization algorithm to compute a single feasible tiling, enumerate all possible tilings within the given region,

Authors' addresses: Chi-Han Peng, KAUST, [chihan.peng@kaust.edu.sa](mailto:chihan.peng@kaust.edu.sa); Helmut Pottmann, TU Wien, [pottmann@geometrie.tuwien.ac.at](mailto:pottmann@geometrie.tuwien.ac.at); Peter Wonka, KAUST, [peter.wonka@kaust.edu.sa](mailto:peter.wonka@kaust.edu.sa).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART107 \$15.00

<https://doi.org/10.1145/3197517.3201306>

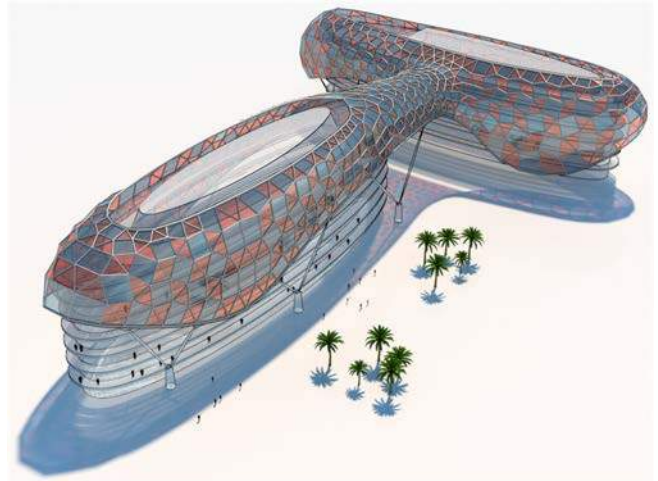


Fig. 1. A triangle-quad hybrid mesh was used to design an alternative outer skin of the Yas Island Marina hotel. The design of regular and near regular triangle-quad patterns on surfaces is a difficult task that is tackled by the computational framework in this paper.

or select a particular tiling by modeling an objective function that includes aesthetic criteria (Section 3.3).

To tackle the design problem in 3D, we combine multiple 2D solutions. We use a graph structure called *patch graph*, to encode the boundaries of individual 2D regular hybrid meshes and cross-patch consistency constraints (Section 4.1). The individual 2D solutions can then be deformed and glued together to yield a 3D design. In some cases, interesting designs can also be generated without deforming the individual 2D solutions, giving rise to an interesting generalization of Lobel frames [Lobel 2004]. Patch graphs can be generated from scratch, by a subdivision scheme, or designed on a given reference surface using our proposed user-interface (Section 4.2).

In Section 5, we showcase various designs enabled by our framework to demonstrate that hybrid meshes are an interesting complement to existing triangle and quad-based meshes. Some advantages of hybrid meshes for fabrication are a finer accuracy of the discretization of angle deficits at a vertex compared to triangle and quad meshes as well as the easier planarization compared to quad meshes. The main contributions of this paper are summarized as follows:

- Introducing the concept of *triangle-quad hybrid meshes*, with which we discovered novel polygonal patterns with a distinct look that cannot be derived from existing triangle/hex or quad meshes.



# Parametric Directional Coding for Precomputed Sound Propagation

NIKUNJ RAGHUVANSHI, Microsoft Research

JOHN SNYDER, Microsoft Research

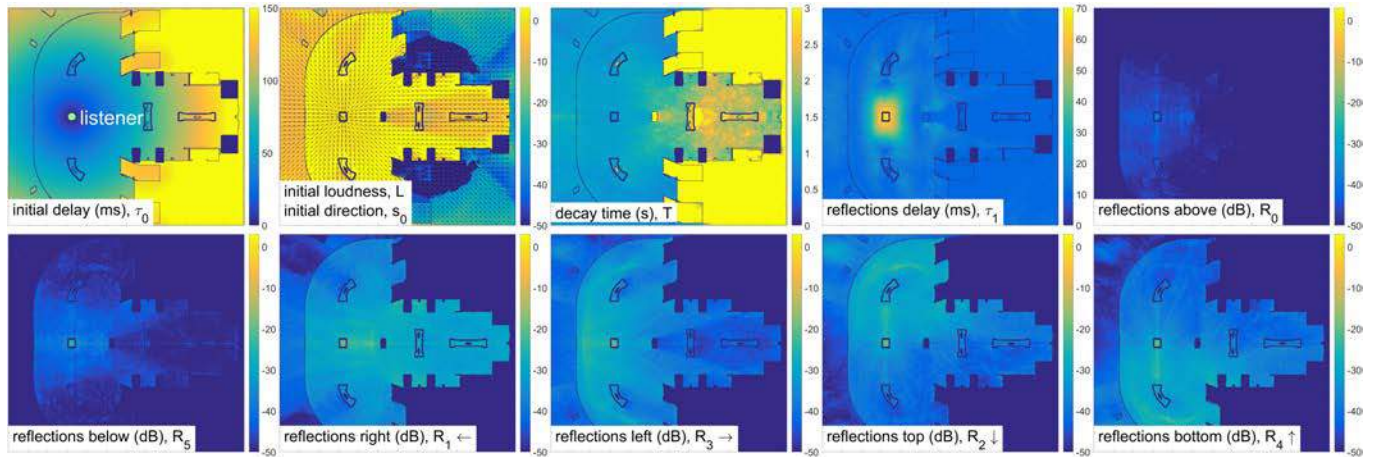


Fig. 1. Parameter fields for HIGHRISE. Source position varies in 3D; one horizontal slice is shown. The listener is held fixed at the green circle.

Convincing audio for games and virtual reality requires modeling directional propagation effects. The initial sound's arrival direction is particularly salient and derives from multiply-diffracted paths in complex scenes. When source and listener straddle occluders, the initial sound and multiply-scattered reverberation stream through gaps and portals, helping the listener navigate. Geometry near the source and/or listener reveals its presence through anisotropic reflections. We propose the first precomputed wave technique to capture such directional effects in general scenes comprising millions of polygons. These effects are formally represented with the 9D directional response function of 3D source and listener location, time, and direction at the listener, making memory use the major concern. We propose a novel parametric encoder that compresses this function within a budget of ~100MB for large scenes, while capturing many salient acoustic effects indoors and outdoors. The encoder is complemented with a lightweight signal processing algorithm whose filtering cost is largely insensitive to the number of sound sources, resulting in an immediately practical system.

CCS Concepts: • **Applied computing** → **Sound and music computing**; • **Computing methodologies** → **Virtual reality**;

Additional Key Words and Phrases: virtual acoustics, sound propagation, spatial audio, directional impulse response, vector intensity, flux density, HRTF, plane wave decomposition, wave equation.

Authors' addresses: Nikunj Raghuvanshi, nikunj@microsoft.com, Microsoft Research; John Snyder, johnsny@microsoft.com, Microsoft Research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART108 \$15.00

<https://doi.org/10.1145/3197517.3201339>

## ACM Reference Format:

Nikunj Raghuvanshi and John Snyder. 2018. Parametric Directional Coding for Precomputed Sound Propagation. *ACM Trans. Graph.* 37, 4, Article 108 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201339>

## 1 INTRODUCTION

Hearing is directional, complementing vision by detecting where sound events occur in our environment. Standing outside a hall, we're able to locate the open door through which streams a crowd's chatter even when the door is obscured or lies behind us. As we walk inside, the auditory scene wraps around us. Meanwhile the initial sound's direction smoothly resolves from the door to each of the speakers, helping us face and navigate to a chosen individual. While reflections envelope the listener indoors, partly open spaces yield anisotropic reflections, reinforcing the visual location of nearby scene geometry. When source and listener are close (e.g. with footsteps), the delay between the initial sound and first reflections can become audible, strengthening the perception of distance to walls.

Our goal is practical modeling and rendering of such directional acoustic effects for games and VR applications in scenes containing millions of polygons. Manual mesh simplification or scene decomposition must be avoided. Many physical effects are perceivable and must be simulated accurately and efficiently as the sound diffracts around obstacles and through portals and scatters many times. Transient effects are critical. Initial sound arriving in the first 1ms, possibly through multiple portals, determines its perceived direction. The directional distribution of later arriving reflections convey additional information about the listener's surroundings.

Ours differs from typical problems in room acoustics [Kuttruff 2000]. We handle real-time movement of both sources and listener, capturing variation in complex scenes without breaking the fraction-of-a-core CPU budget typical for games. Effects should be robust even in occluded cases where conventional path tracing demands

# Toward Wave-based Sound Synthesis for Computer Animation

JUI-HSIEN WANG, Stanford University

ANTE QU, Stanford University

TIMOTHY R. LANGLOIS, Adobe Research

DOUG L. JAMES, Stanford University

We explore an integrated approach to sound generation that supports a wide variety of physics-based simulation models and computer-animated phenomena. Targeting high-quality offline sound synthesis, we seek to resolve animation-driven sound radiation with near-field scattering and diffraction effects. The core of our approach is a sharp-interface finite-difference time-domain (FDTD) wavesolver, with a series of supporting algorithms to handle rapidly deforming and vibrating embedded interfaces arising in physics-based animation sound. Once the solver rasterizes these interfaces, it must evaluate acceleration boundary conditions (BCs) that involve model- and phenomena-specific computations. We introduce *acoustic shaders* as a mechanism to abstract away these complexities, and describe a variety of implementations for computer animation: near-rigid objects with ringing and acceleration noise, deformable (finite element) models such as thin shells, bubble-based water, and virtual characters. Since time-domain wave synthesis is expensive, we only simulate pressure waves in a small region about each sound source, then estimate a far-field pressure signal. To further improve scalability beyond multi-threading, we propose a fully *time-parallel sound synthesis* method that is demonstrated on commodity cloud computing resources. In addition to presenting results for multiple animation phenomena (water, rigid, shells, kinematic deformers, etc.) we also propose 3D automatic dialogue replacement (3DADR) for virtual characters so that pre-recorded dialogue can include character movement, and near-field shadowing and scattering sound effects.

CCS Concepts: • **Computing methodologies** → *Physical simulation*;

Additional Key Words and Phrases: Computer animation, sound synthesis, finite-difference time-domain method, acoustics

## ACM Reference Format:

Jui-Hsien Wang, Ante Qu, Timothy R. Langlois, and Doug L. James. 2018. Toward Wave-based Sound Synthesis for Computer Animation. *ACM Trans. Graph.* 37, 4, Article 109 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201318>

## 1 INTRODUCTION

Recent advances in physics-based sound synthesis have led to improved sound-generation techniques for computer-animated phenomena, including water, rigid bodies, deformable models like rods

and shells, fire, and brittle fracture. However, unlike visual rendering techniques which are slow but capable of generating very high-quality, general-purpose results during parallel offline rendering, current methods for sound synthesis suffer from several fundamental deficiencies that limit high-quality renderings. First, because of the high space-time complexity of sound modeling and complicated acoustic phenomena, many modeling approximations (linearizations, simplified radiation, etc.) and algorithmic optimizations (precomputation, data-driven methods, etc.) are typically made for tractability and to improve performance, but can limit sound quality and the range of physical phenomena that can be considered. Second, the proliferation of real-time, object-level and/or phenomena-specific sound models has led to a lack of integrated wave modeling, wherein it is difficult to combine models and have them support inter-model acoustic interactions, such as shadowing and scattering. The end result is that there is no practical general-purpose solution (regardless of performance) for practitioners to author general animations with integrated sound synthesis. In contrast, we seek a general rendering approach where, given an animation, say of a crash cymbal being hit (see Figure 1), one can just “hit the render button and wait,” to get the high-quality result.

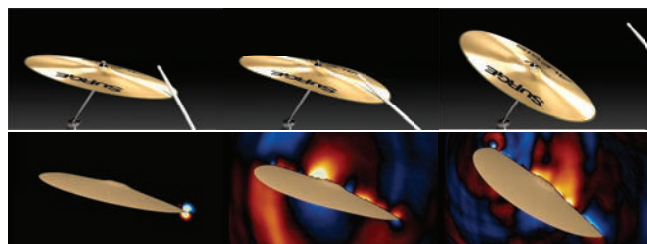


Fig. 1. **Crash! Sound waves radiate off a visibly deforming crash cymbal:** Our sharp-interface finite-difference time-domain (FDTD) wave solver can synthesize sound sources with rapidly deforming interfaces, such as this thin-shell model, or even water, using a single integrated approach.

In this paper, we explore offline wave-based sound synthesis techniques that can generate high-quality animation-sound content for general-purpose dynamic scenes and multi-physics sound sources. The general-purpose integrated approach is inherently extremely slow, but enables “what if” exploration of the rich range of sounds possible from complex animated phenomena (such as splashing water), and to better understand future challenges and needs for animation sound. We humbly liken our approach to early “what if” digital image synthesis efforts three decades ago, such as the “one frame movie” generated in the ‘80s by the aspiring REYES (“Renders Everything You Ever Saw”) 3D rendering system [Cook et al. 1987] that “demonstrated, at least theoretically, the possibility of creating

Authors’ addresses: Jui-Hsien Wang, Stanford University, [jui-hsien.wang@stanford.edu](mailto:jui-hsien.wang@stanford.edu); Ante Qu, Stanford University, [antegu@cs.stanford.edu](mailto:antegu@cs.stanford.edu); Timothy R. Langlois, Adobe Research, [tlangloi@adobe.com](mailto:tlangloi@adobe.com); Doug L. James, Stanford University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART109 \$15.00

<https://doi.org/10.1145/3197517.3201318>

# Multi-Scale Simulation of Nonlinear Thin-Shell Sound with Wave Turbulence

GABRIEL CIRIO, Inria, Université Côte d’Azur and Columbia University

ANTE QU, Stanford University

GEORGE DRETTAKIS, Inria and Université Côte d’Azur

EITAN GRINSPUN and CHANGXI ZHENG, Columbia University

Thin shells — solids that are thin in one dimension compared to the other two — often emit rich nonlinear sounds when struck. Strong excitations can even cause chaotic thin-shell vibrations, producing sounds whose energy spectrum diffuses from low to high frequencies over time — a phenomenon known as wave turbulence. It is all these nonlinearities that grant shells such as cymbals and gongs their characteristic “glinting” sound. Yet, simulation models that efficiently capture these sound effects remain elusive.

We propose a physically based, multi-scale reduced simulation method to synthesize nonlinear thin-shell sounds. We first split nonlinear vibrations into two scales, with a small low-frequency part simulated in a fully nonlinear way, and a high-frequency part containing many more modes approximated through time-varying linearization. This allows us to capture interesting nonlinearities in the shells’ deformation, tens of times faster than previous approaches. Furthermore, we propose a method that enriches simulated sounds with wave turbulent sound details through a phenomenological diffusion model in the frequency domain, and thereby sidestep the expensive simulation of chaotic high-frequency dynamics. We show several examples of our simulations, illustrating the efficiency and realism of our model.

CCS Concepts: • **Computing methodologies** → **Physical simulation**; • **Applied computing** → Sound and music computing;

Additional Key Words and Phrases: Sound synthesis, thin shells, reduced simulation, wave turbulence

## ACM Reference Format:

Gabriel Cirio, Ante Qu, George Drettakis, Eitan Grinspun, and Changxi Zheng. 2018. Multi-Scale Simulation of Nonlinear Thin-Shell Sound with Wave Turbulence. *ACM Trans. Graph.* 37, 4, Article 110 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201361>

## 1 INTRODUCTION

Thin shells produce rich and complex sounds: from containers such as trash cans and plastic bottles, to musical instruments like cymbals or gongs. Once excited, vibrational displacements as small as the thickness trigger highly nonlinear dynamics, such as time-varying amplitudes of the spectrum harmonics, energy exchange between vibrational modes, and ultimately chaos and turbulence [Chaigne et al. 2005]. As a result, thin-shell sound has time-varying details across a wide frequency spectrum. In this paper, we present a method

Authors’ addresses: Gabriel Cirio, Inria, Université Côte d’Azur and Columbia University; Ante Qu, Stanford University; George Drettakis, Inria and Université Côte d’Azur; Eitan Grinspun; Changxi Zheng, Columbia University.

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2018/8-ART110 \$15.00

<https://doi.org/10.1145/3197517.3201361>

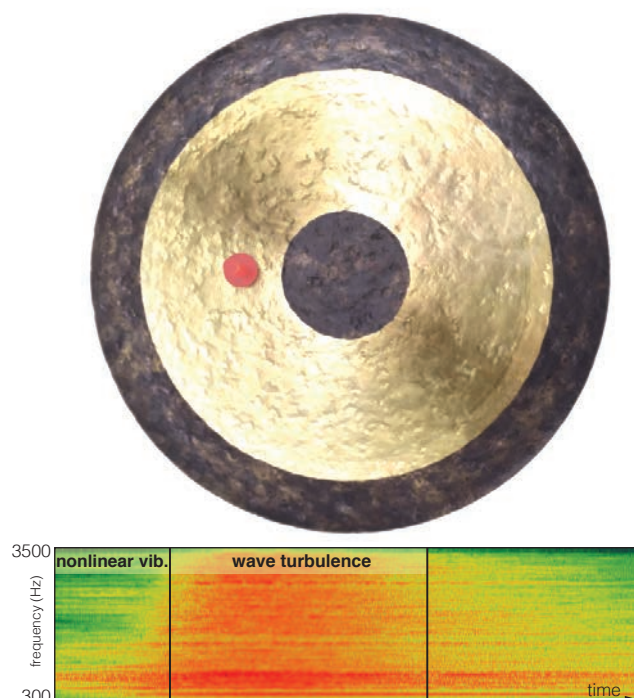


Fig. 1. **Splashy Tamtam.** A Tamtam gong produces rich nonlinear sounds. When it is struck hard (e.g., at the position indicated by the red dot), its vibrational energy quickly cascades from low to high frequencies — a phenomenon known as *wave turbulence*, and produces rich sound details in a wide spectrum. Our method is able to efficiently simulate nonlinear thin-shell sounds, and enrich them with a fast phenomenological model of wave turbulence. The nonlinear sound and the emergence of wave turbulence are shown in the spectrogram on the bottom and the supplemental video.

that efficiently captures these challenging nonlinear sound effects, including wave turbulence in particular.

Recent advances in physics-based sound models have enabled the synthesis of nonlinear sounds from thin shells, often at a prohibitive simulation cost [O’Brien et al. 2001]. The Harmonic Shells model [Chadwick et al. 2009] significantly reduces the runtime simulation cost, but requires a heavy precomputation process that learns a thin-shell cubature scheme from full simulation data. While the resulting sounds capture certain nonlinearities such as pitch glide, other important nonlinear effects such as the turbulence-like energy cascade are still missing.



# Scene-Aware Audio for 360° Videos

DINGZEYU LI, Columbia University

TIMOTHY R. LANGLOIS, Adobe Research

CHANGXI ZHENG, Columbia University

Although 360° cameras ease the capture of panoramic footage, it remains challenging to add realistic 360° audio that blends into the captured scene and is synchronized with the camera motion. We present a method for adding scene-aware spatial audio to 360° videos in typical indoor scenes, using only a conventional mono-channel microphone and a speaker. We observe that the late reverberation of a room's impulse response is usually diffuse spatially and directionally. Exploiting this fact, we propose a method that synthesizes the directional impulse response between any source and listening locations by combining a synthesized early reverberation part and a measured late reverberation tail. The early reverberation is simulated using a geometric acoustic simulation and then enhanced using a frequency modulation method to capture room resonances. The late reverberation is extracted from a recorded impulse response, with a carefully chosen time duration that separates out the late reverberation from the early reverberation. In our validations, we show that our synthesized spatial audio matches closely with recordings using ambisonic microphones. Lastly, we demonstrate the strength of our method in several applications.

CCS Concepts: • **Computing methodologies** → **Physical Simulation**; **Virtual reality**; **Mixed / Augmented reality**;

Additional Key Words and Phrases: 360° videos, ambisonic audio

## ACM Reference Format:

Dingzeyu Li, Timothy R. Langlois, and Changxi Zheng. 2018. Scene-Aware Audio for 360° Videos. *ACM Trans. Graph.* 37, 4, Article 111 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201391>

## 1 INTRODUCTION

The ecosystem of 360° video is flourishing. Devices such as the Samsung Gear 360 and the Ricoh Theta have facilitated 360° video capture; software such as Adobe Premiere Pro has included features for editing 360° panoramic footage; and online platforms such as Youtube and Facebook have promoted easy sharing and viewing of 360° content. With these technological advances, video creators now have a whole new set of tools for creating immersive visual experiences. Yet, the creation of their auditory accompaniment, the immersive audio, is not as easy. Immersive 360° videos are noticeably lacking immersive scene-aware 360° audio.

Toward filling this gap, we propose a method that enables 360° video creators to easily add spatial audio from specified sound

Authors' addresses: Dingzeyu Li, Columbia University, [dli@cs.columbia.edu](mailto:dli@cs.columbia.edu); Timothy R. Langlois, Adobe Research, [tlangloi@adobe.com](mailto:tlangloi@adobe.com); Changxi Zheng, Columbia University, [czx@cs.columbia.edu](mailto:czx@cs.columbia.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART111 \$15.00

<https://doi.org/10.1145/3197517.3201391>

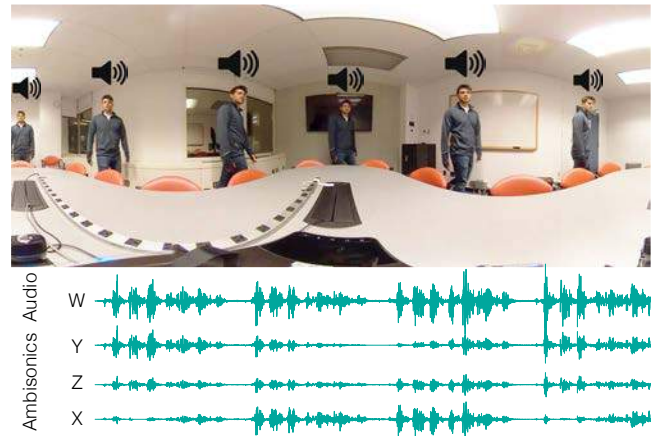


Fig. 1. **360° audiovisual capture.** Our method enables video creators to add ambisonic audio (bottom) in a 360° video of a indoor scene (top). When viewers watch the video and change the camera angle, they hear the binaural audio consistent with the current viewing angle. Our method has no restriction on the input audio. In the case shown here, the input audio is the person's speech captured using a conventional mono-channel microphone collocated with the 360° camera, and our method converts the mono-channel input audio into a spatial audio in standard ambisonic format. The waveforms show a first-order ambisonic output (four channels), although our method supports an arbitrary order of ambisonics.

sources in a typical indoor scene, such as the conference room shown in Figure 1. Our method consists of two stages. We first record a single acoustic impulse response in a room using a readily available mono-channel microphone and a simple setup. Then, provided any 360° footage captured in the same environment and a piece of source audio, our method outputs the 360° video with an accompanying ambisonics spatial soundtrack. The resulting soundfield captures the spatial sound effects at the camera location, even if the camera is dynamic, as if the input audio is emitted from a user-specified sound source in the environment. Our method has no restriction on the input audio: it could be artificially synthesized, recorded in an anechoic chamber, or recorded in the same scene simply using a conventional mono-channel microphone (Figure 1). The generated ambisonic audio can be directly played back in realtime by a spatial audio player.

A conventional microphone and a sound source are the only requirements our method, in addition to a 360° camera. This contrasts starkly with the current approach of capturing spatial audio, which requires the use of a soundfield ambisonic microphone, which uses a microphone array consisting of multiple carefully positioned mono-channel microphones to record the spatial sound field. These devices are generally expensive, and currently very few 360° cameras have an integrated ambisonic microphone. When designing sound for traditional media, audio from each source is processed to

# Looking to Listen at the Cocktail Party: A Speaker-Independent Audio-Visual Model for Speech Separation

ARIEL EPHRAT, Google Research and The Hebrew University of Jerusalem, Israel

INBAR MOSSERI, Google Research

ORAN LANG, Google Research

TALI DEKEL, Google Research

KEVIN WILSON, Google Research

AVINATAN HASSIDIM, Google Research

WILLIAM T. FREEMAN, Google Research

MICHAEL RUBINSTEIN, Google Research

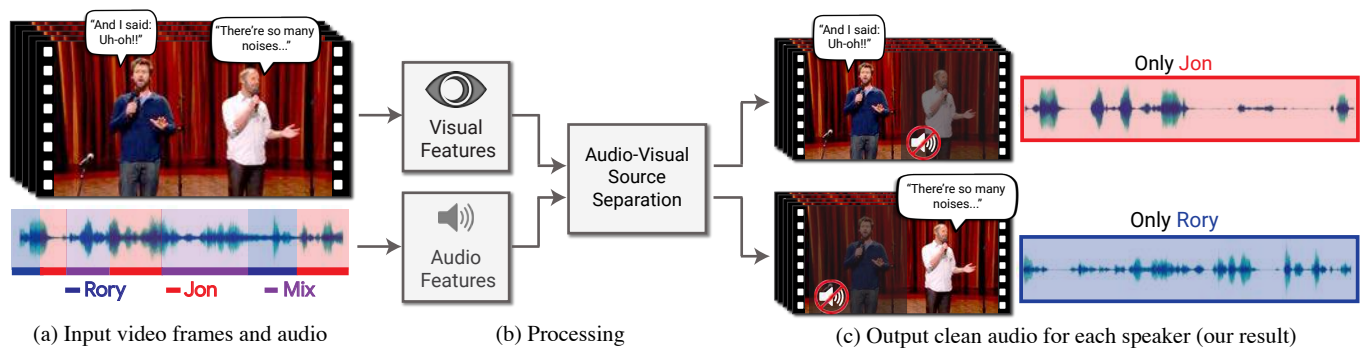


Fig. 1. We present a model for isolating and enhancing the speech of desired speakers in a video. (a) The input is a video (frames + audio track) with one or more people speaking, where the speech of interest is interfered by other speakers and/or background noise. (b) Both audio and visual features are extracted and fed into a joint audio-visual speech separation model. The output is a decomposition of the input audio track into clean speech tracks, one for each person detected in the video (c). This allows us to then compose videos where speech of specific people is enhanced while all other sound is suppressed. Our model was trained using thousands of hours of video segments from our new dataset, AVSPEECH. The “Stand-Up” video (a) is courtesy of Team Coco.

We present a joint audio-visual model for isolating a single speech signal from a mixture of sounds such as other speakers and background noise. Solving this task using only audio as input is extremely challenging and does not provide an association of the separated speech signals with speakers in the video. In this paper, we present a deep network-based model that incorporates both visual and auditory signals to solve this task. The visual features are used to “focus” the audio on desired speakers in a scene and to improve the speech separation quality. To train our joint audio-visual model, we introduce AVSPEECH, a new dataset comprised of thousands of hours of video segments from the Web. We demonstrate the applicability of our method to classic speech separation tasks, as well as real-world scenarios involving heated interviews, noisy bars, and screaming children, only requiring the user to specify the face of the person in the video whose speech they want to isolate. Our method shows clear advantage over state-of-the-art audio-only speech separation in cases of mixed speech. In addition, our model, which is speaker-independent (trained once, applicable to any speaker), produces better results than recent audio-visual speech separation

methods that are speaker-dependent (require training a separate model for each speaker of interest).

Additional Key Words and Phrases: Audio-Visual, Source Separation, Speech Enhancement, Deep Learning, CNN, BLSTM

## ACM Reference Format:

Ariel Ephrat, Inbar Mosseri, Oran Lang, Tali Dekel, Kevin Wilson, Avinatan Hassidim, William T. Freeman, and Michael Rubinstein. 2018. Looking to Listen at the Cocktail Party: A Speaker-Independent Audio-Visual Model for Speech Separation. *ACM Trans. Graph.* 37, 4, Article 112 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201357>

## 1 INTRODUCTION

Humans are remarkably capable of focusing their auditory attention on a single sound source within a noisy environment, while de-emphasizing (“muting”) all other voices and sounds. The way neural systems achieve this feat, which is known as the *cocktail party effect* [Cherry 1953], remains unclear. However, research has shown that viewing a speaker’s face enhances a person’s capacity to resolve perceptual ambiguity in a noisy environment [Golumbic et al. 2013; Ma et al. 2009]. In this paper we achieve this ability computationally.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART112

<https://doi.org/10.1145/3197517.3201357>

The first author performed this work as an intern at Google.

# Single-Photon 3D Imaging with Deep Sensor Fusion

DAVID B. LINDELL, Stanford University, USA  
MATTHEW O'TOOLE, Stanford University, USA  
GORDON WETZSTEIN, Stanford University, USA

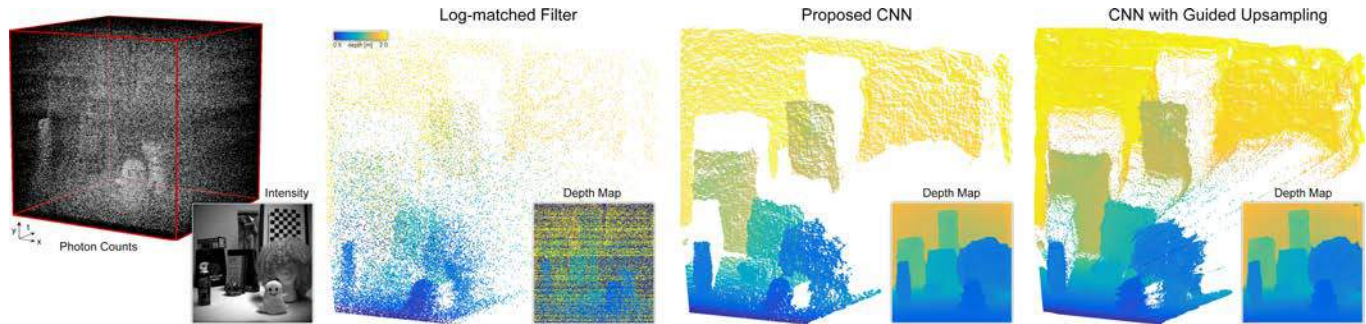


Fig. 1. Single-photon 3D imaging systems measure a spatio-temporal volume containing photon counts (left) that include ambient light, noise, and photons emitted by a pulsed laser into the scene and reflected back to the detector. Conventional depth estimation techniques, such as log-matched filtering (center left), estimate a depth map from these counts. However, depth estimation is a non-convex and challenging problem, especially for extremely low photon counts observed in fast or long-range 3D imaging systems. We introduce a data-driven approach to solve this depth estimation problem and explore deep sensor fusion approaches that use an intensity image of the scene to optimize the robustness (center right) and resolution (right) of the depth estimation.

Sensors which capture 3D scene information provide useful data for tasks in vehicle navigation, gesture recognition, human pose estimation, and geometric reconstruction. Active illumination time-of-flight sensors in particular have become widely used to estimate a 3D representation of a scene. However, the maximum range, density of acquired spatial samples, and overall acquisition time of these sensors is fundamentally limited by the minimum signal required to estimate depth reliably. In this paper, we propose a data-driven method for photon-efficient 3D imaging which leverages sensor fusion and computational reconstruction to rapidly and robustly estimate a dense depth map from low photon counts. Our sensor fusion approach uses measurements of single photon arrival times from a low-resolution single-photon detector array and an intensity image from a conventional high-resolution camera. Using a multi-scale deep convolutional network, we jointly process the raw measurements from both sensors and output a high-resolution depth map. To demonstrate the efficacy of our approach, we implement a hardware prototype and show results using captured data. At low signal-to-background levels, our depth reconstruction algorithm with sensor fusion outperforms other methods for depth estimation from noisy measurements of photon arrival times.

Authors' addresses: David B. Lindell, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, lindell@stanford.edu; Matthew O'Toole, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, matthew.otoole@gmail.com; Gordon Wetzstein, Department of Electrical Engineering, Stanford University, Stanford, CA, 94305, USA, gordon.wetzstein@stanford.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://doi.org/10.1145/3197517.3201316).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART113 \$15.00

<https://doi.org/10.1145/3197517.3201316>

CCS Concepts: • **Computing methodologies** → **3D imaging**; *Neural networks*; *Computational photography*;

Additional Key Words and Phrases: computational photography, single-photon imaging

## ACM Reference Format:

David B. Lindell, Matthew O'Toole, and Gordon Wetzstein. 2018. Single-Photon 3D Imaging with Deep Sensor Fusion. *ACM Trans. Graph.* 37, 4, Article 113 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201316>

## 1 INTRODUCTION

3D imaging systems provide scene information crucial for a diverse range of applications, including autonomous vehicles, robotic vision, 3D object modeling, gesture recognition, pose tracking, scene understanding, segmentation, localization and mapping, and remote sensing. Depth sensors using the time-of-flight (ToF) principle are among the most popular choices with pulsed systems being most suitable for long-range outdoor applications [Horaud et al. 2016; Koskinen et al. 1992].

Pulsed ToF or light detection and ranging (LIDAR) systems often employ a co-axially aligned laser diode and single-photon detector. The laser emits a short pulse, usually on the order of hundreds of picoseconds or a few nanoseconds, and the detector timestamps the arrival of photons reflected back by the scene. Thus, for a single emitted pulse or a sequence of pulses, the detector records a temporal histogram of photon counts. A 3D spatio-temporal volume containing raw photon counts is then acquired by either scanning the laser-detector pair, using an array of these pairs, or using a hybrid approach of scanning and arraying.

Robustly estimating depth from raw photon counts acquired with pulsed ToF systems in low-flux scenarios is a major challenge [McCarthy et al. 2013; Pawlikowska et al. 2017]. When only a few signal



# End-to-end Optimization of Optics and Image Processing for Achromatic Extended Depth of Field and Super-resolution Imaging

VINCENT SITZMANN\*, Stanford University, USA

STEVEN DIAMOND\*, Stanford University, USA

YIFAN PENG\*, The University of British Columbia, Canada and Stanford University, USA

XIONG DUN, King Abdullah University of Science and Technology, Saudi Arabia

STEPHEN BOYD, Stanford University, USA

WOLFGANG HEIDRICH, King Abdullah University of Science and Technology, Saudi Arabia

FELIX HEIDE, Stanford University, USA

GORDON WETZSTEIN, Stanford University, USA



Fig. 1. One of the applications of the proposed end-to-end computational camera design paradigm is achromatic extended depth of field. When capturing an image with a regular singlet lens (top left), out-of-focus regions are blurry and chromatic aberrations further degrade the image quality. With our framework, we optimize the profile of a refractive optical element that achieves both depth and chromatic invariance. This element is fabricated using diamond turning (right) or using photolithography. After processing an image recorded with this optical element using a simple Wiener deconvolution, we obtain an all-in-focus image with little chromatic aberrations (top center). Point spread functions for both the regular lens and the optimized optical element are shown in the bottom. In this paper, we explore several applications that demonstrate the efficacy of our novel approach to domain-specific computational camera design.

Authors' addresses: Vincent Sitzmann\*, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, [sitzmann@cs.stanford.edu](mailto:sitzmann@cs.stanford.edu); Steven Diamond\*, Stanford University, Department of Computer Science, Stanford, CA, 94305, USA, [diamond@cs.stanford.edu](mailto:diamond@cs.stanford.edu); Yifan Peng\*, The University of British Columbia, Department of Computer Science, Vancouver, V6T 1Z4, Canada, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, [evanpeng@cs.ubc.ca](mailto:evanpeng@cs.ubc.ca); Xiong Dun, King Abdullah University of Science and Technology, Visual Computing Center, Thuwal, 23955, Saudi Arabia, [xiong.dun@kaust.edu.sa](mailto:xiong.dun@kaust.edu.sa); Stephen Boyd, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, [boyd@stanford.edu](mailto:boyd@stanford.edu); Wolfgang Heidrich, King Abdullah University of Science and Technology, Visual Computing Center, Thuwal, 23955, Saudi Arabia, [wolfgang.heidrich@kaust.edu.sa](mailto:wolfgang.heidrich@kaust.edu.sa); Felix Heide, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, [fheide@stanford.edu](mailto:fheide@stanford.edu); Gordon Wetzstein, Stanford University, Department of Electrical Engineering, Stanford, CA, 94305, USA, [gordon.wetzstein@stanford.edu](mailto:gordon.wetzstein@stanford.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

In typical cameras the optical system is designed first; once it is fixed, the parameters in the image processing algorithm are tuned to get good image reproduction. In contrast to this sequential design approach, we consider joint optimization of an optical system (for example, the physical shape of the lens) together with the parameters of the reconstruction algorithm. We build a fully-differentiable simulation model that maps the true source image to the reconstructed one. The model includes diffractive light propagation, depth and wavelength-dependent effects, noise and nonlinearities, and the image post-processing. We jointly optimize the optical parameters and the image processing algorithm parameters so as to minimize the deviation between the true and reconstructed image, over a large set of images. We implement our joint optimization method using autodifferentiation to efficiently compute parameter gradients in a stochastic optimization algorithm. We demonstrate the efficacy of this approach by applying it to achromatic extended depth of field and snapshot super-resolution imaging.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
0730-0301/2018/8-ART114 \$15.00  
<https://doi.org/10.1145/3197517.3201333>

# Megapixel Adaptive Optics: Towards Correcting Large-scale Distortions in Computational Cameras

CONGLI WANG, QIANG FU, XIONG DUN, and WOLFGANG HEIDRICH,  
King Abdullah University of Science and Technology, Saudi Arabia

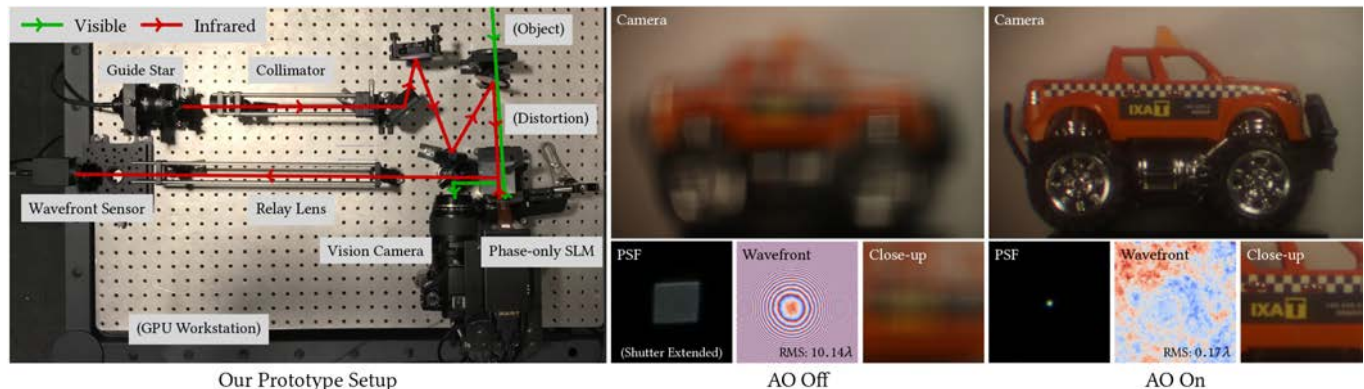


Fig. 1. We propose to use a high resolution adaptive optics (AO) technique to correct for large distortions in photography and vision applications. AO relies on the combination of a wavefront sensor and a phase modulator to measure and optically correct for incoming phase distortions. Based on a previously proposed high resolution wavefront sensor, our proposed AO prototype achieves unprecedented AO sensing and correcting simultaneous high spatial resolution for large phase distortions. An example for vision deblurring is shown on the right where a defocus distortion is being well compensated.

Adaptive optics has become a valuable tool for correcting minor optical aberrations in applications such as astronomy and microscopy. However, due to the limited resolution of both the wavefront sensing and the wavefront correction hardware, it has so far not been feasible to use adaptive optics for correcting large-scale waveform deformations that occur naturally in regular photography and other imaging applications.

In this work, we demonstrate an adaptive optics system for regular cameras. We achieve a significant improvement in focus for large wavefront distortions by improving upon a recently developed high resolution coded wavefront sensor, and combining it with a spatial phase modulator to create a megapixel adaptive optics system with unprecedented capability to sense and correct large distortions.

CCS Concepts: • **Computing methodologies** → **Computational photography**;

Additional Key Words and Phrases: Adaptive Optics, Wavefront Sensing, Computational Cameras

## ACM Reference Format:

Congli Wang, Qiang Fu, Xiong Dun, and Wolfgang Heidrich. 2018. Megapixel Adaptive Optics: Towards Correcting Large-scale Distortions in Computational Cameras. *ACM Trans. Graph.* 37, 4, Article 115 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201299>

Authors' addresses: C. Wang, [congli.wang@kaust.edu.sa](mailto:congli.wang@kaust.edu.sa); Q. Fu, [qiang.fu@kaust.edu.sa](mailto:qiang.fu@kaust.edu.sa); X. Dun, [xiong.dun@kaust.edu.sa](mailto:xiong.dun@kaust.edu.sa); W. Heidrich, [wolfgang.heidrich@kaust.edu.sa](mailto:wolfgang.heidrich@kaust.edu.sa), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia, 23955-6900.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART115

<https://doi.org/10.1145/3197517.3201299>

## 1 INTRODUCTION

Adaptive Optics (AO) systems are highly effective in correcting dynamic aberrations in applications such as astronomy and microscopy [Beckers 1993; Marsh et al. 2003]. In these settings, the aberrations are typically small (e.g. atmospheric distortions in the case of telescopes), albeit significant in the case of otherwise diffraction limited optics.

Adaptive optics systems are comprised of two major components: a wavefront sensor, which measures the shape of the wavefront of light across the aperture of the optical system, and a phase modulator, which corrects the distorted wavefront to make it planar. In traditional AO systems, the wavefront sensor is usually a Shack-Hartmann [Shack and Platt 1971] or pyramid sensor [Ragazzoni 1996], while the phase modulation is achieved by tilting mirrors [Bifano et al. 1999; Perreault et al. 2002] or deforming a reflective membrane with pistons [Zhu et al. 1999]. Both the sensing and the modulation components in these systems suffer from low *spatial resolution* (e.g. at most hundreds to thousands of measurement points and actuators) as well as small *range* that limits the amplitude of the distortion. As a result, most AO systems can only correct for distortions corresponding to a few low-order Zernike polynomials and low amplitudes. Multi-conjugate AO systems [Beckers 1988; Diolaiti et al. 2001] are one possible solution to offer high resolution and large compensation amplitudes, but at the cost of additional wavefront sensors and correctors that have to be carefully aligned and calibrated to one another. This significantly increases the cost and overall system complexity compared to the single sensor / single corrector systems that we consider in our work.

# What Characterizes Personalities of Graphic Designs?

NANXUAN ZHAO, City University of Hong Kong

YING CAO, City University of Hong Kong

RYNISON W.H. LAU, City University of Hong Kong

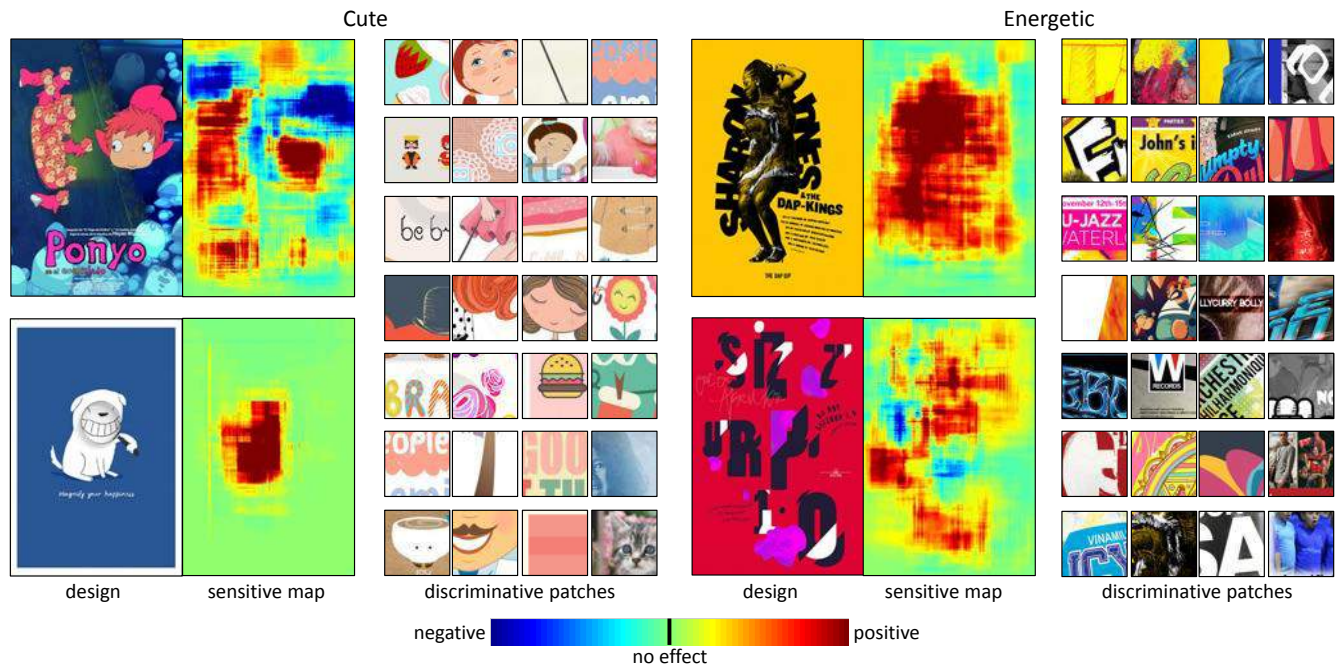


Fig. 1. Our model can automatically discover what characterizes the personality of a graphic design. For each of the two personality labels (“Cute” and “Energetic”) shown above, we show two design examples together with their corresponding personality-based sensitive maps. This sensitive map measures the contribution of each local region of a design example to a particular personality, with red indicating positive effect while blue indicating negative effect. For each personality, we also show a set of discriminative patches that contribute most to the personality. These patches are mined from a collection of designs that are ranked top by our model based on the personality. ©Hayao Miyazaki, I Love Doodle, Scott Williams and Andrea Dell’Anna.

Graphic designers often manipulate the overall look and feel of their designs to convey certain personalities (e.g., cute, mysterious and romantic) to impress potential audiences and achieve business goals. However, understanding the factors that determine the personality of a design is challenging, as a graphic design is often a result of thousands of decisions on numerous factors, such as font, color, image, and layout. In this paper, we aim to answer the question of what characterizes the personality of a graphic design. To

Ying Cao is the corresponding author. This work was led by Rynson Lau. Authors’ addresses: Nanxuan Zhao, Department of Computer Science, City University of Hong Kong, nanxuzhao2-c@my.cityu.edu.hk; Ying Cao, Department of Computer Science, City University of Hong Kong, caoying59@gmail.com; Rynson W.H. Lau, Department of Computer Science, City University of Hong Kong, rynson.lau@cityu.edu.hk.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART116 \$15.00

<https://doi.org/10.1145/3197517.3201355>

this end, we propose a deep learning framework for exploring the effects of various design factors on the perceived personalities of graphic designs. Our framework learns a convolutional neural network (called *personality scoring network*) to estimate the personality scores of graphic designs by ranking the crawled web data. Our personality scoring network automatically learns a visual representation that captures the semantics necessary to predict graphic design personality. With our personality scoring network, we systematically and quantitatively investigate how various design factors (e.g., color, font, and layout) affect design personality across different scales (from pixels, regions to elements). We also demonstrate a number of practical application scenarios of our network, including element-level design suggestion and example-based personality transfer.

CCS Concepts: • **Computing methodologies** → **Perception**; Neural networks;

Additional Key Words and Phrases: Graphic design, Personality, Deep learning

## ACM Reference Format:

Nanxuan Zhao, Ying Cao, and Rynson W.H. Lau. 2018. What Characterizes Personalities of Graphic Designs?. *ACM Trans. Graph.* 37, 4, Article 116 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201355>



# Scale-aware Black-and-White Abstraction of 3D Shapes

YOU-EN LIN, National Tsing Hua University  
YONG-LIANG YANG, University of Bath  
HUNG-KUO CHU, National Tsing Hua University

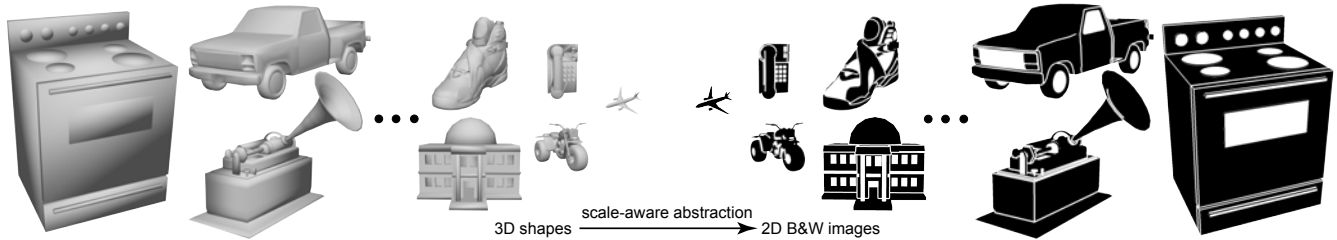


Fig. 1. Proposed computational framework for automatic generation of 2D black-and-white abstractions from 3D shapes which preserves geometric and structural properties at different scales. Scales of abstraction results in both paper and supplementary material were adaptively selected such that feature lines are recognizable and consistent over different scales.

Flat design is a modern style of graphics design that minimizes the number of design attributes required to convey 3D shapes. This approach suits design contexts requiring simplicity and efficiency, such as mobile computing devices. This ‘less-is-more’ design inspiration has posed significant challenges in practice since it selects from a restricted range of design elements (e.g., color and resolution) to represent complex shapes. In this work, we investigate a means of computationally generating a specialized 2D flat representation - image formed by black-and-white patches - from 3D shapes. We present a novel framework that automatically abstracts 3D man-made shapes into 2D binary images at multiple scales. Based on a set of identified design principles related to the inference of geometry and structure, our framework jointly analyzes the input 3D shape and its counterpart 2D representation, followed by executing a carefully devised layout optimization algorithm. The robustness and effectiveness of our method are demonstrated by testing it on a wide variety of man-made shapes and comparing the results with baseline methods via a pilot user study. We further present two practical applications that are likely to benefit from our work.

CCS Concepts: • **Computing methodologies** → **Computer graphics**;

Additional Key Words and Phrases: black-and-white image, scale-aware abstraction, joint 2D/3D analysis, layout optimization

## ACM Reference Format:

You-En Lin, Yong-Liang Yang, and Hung-Kuo Chu. 2018. Scale-aware Black-and-White Abstraction of 3D Shapes. *ACM Trans. Graph.* 37, 4, Article 1 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201372>

Authors’ addresses: You-En Lin, Department of Computer Science, National Tsing Hua University, [unlin@livemail.tw](mailto:unlin@livemail.tw); Yong-Liang Yang, Department of Computer Science, University of Bath, [y.yang@cs.bath.ac.uk](mailto:y.yang@cs.bath.ac.uk); Hung-Kuo Chu, Department of Computer Science, National Tsing Hua University, [hkchu@cs.nthu.edu.tw](mailto:hkchu@cs.nthu.edu.tw).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART1 \$15.00

<https://doi.org/10.1145/3197517.3201372>

## 1 INTRODUCTION

With the recent development of mobile computing technologies, a fascinating design style called *flat design* has gained ever-increasing popularity. In contrast to ‘skeuomorphic’ design, which aims for a high degree of realism using ornamental design attributes (e.g., shadows, gradients, textures, or decorations, etc.), flat design relies on the minimum number of stylistic elements required for the illusion of three dimensions [Page 2014]. The resulting design outcomes hold a number of advantages, such as improved readability and legibility, easy compatibility of design style, and less network and memory cost, etc. These advantages serve a range of design contexts applicable to mobile apps, PC desktops, and webpages.

Among different styles of flat design, a specialization that abstracts 3D shapes into 2D binary images is particularly interesting. It applies the minimum set of colors (i.e., black and white) necessary for 2D design to create a concise abstraction of its 3D counterpart (see Figure 2). Due to the constraints on colors and the reduction of dimensions from 3D to 2D, how to preserve the defining characteristics of the 3D geometry and allow visual perception of the underlying 3D shapes become extremely important.

Various intriguing design elements have been used by designers to better infer 3D geometry. For example, as shown in Figure 2, to reflect shape occlusion and facilitate depth perception, designers often introduce a *halo-like gap region* on occluded sections. Designers add



Fig. 2. Example designs of 2D black-and-white abstractions of 3D shapes created by designers. Design elements such as halo-like gap region (red box), hollow line (blue box) and patch (green box) are used to better infer 3D geometry. Image courtesy of <https://iconmonstr.com/>.

# Perception-Driven Semi-Structured Boundary Vectorization

SHAYAN HOSHYARI, University of British Columbia

EDOARDO ALBERTO DOMINICI, University of British Columbia

ALLA SHEFFER, University of British Columbia

NATHAN CARR, Adobe

ZHAOWEN WANG, Adobe

DUYGU CEYLAN, Adobe

I-CHAO SHEN, National Taiwan University

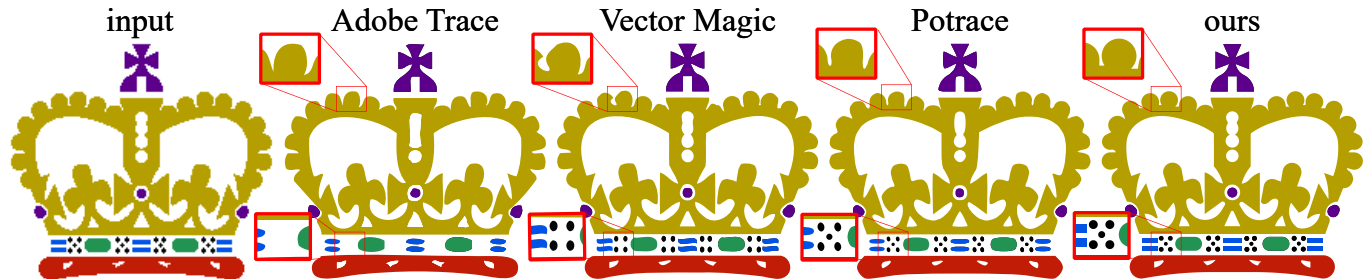


Fig. 1. Vectorizing artist-drawn semi-structured raster images (left) using existing methods (center) results in visible artifacts (highlighted in zoomed-in insets). Our perception-driven vectorization output (right) is consistent with viewer expectations. Please zoom in online to see details. Input image © aves – stock.adobe.com.

Artist-drawn images with distinctly colored, piecewise continuous boundaries, which we refer to as *semi-structured imagery*, are very common in online raster databases and typically allow for a perceptually unambiguous mental vector interpretation. Yet, perhaps surprisingly, existing vectorization algorithms frequently fail to generate these viewer-expected interpretations on such imagery. In particular, the vectorized region boundaries they produce frequently diverge from those anticipated by viewers. We propose a new approach to region boundary vectorization that targets semi-structured inputs and leverages observations about human perception of shapes to generate vector images consistent with viewer expectations. When viewing raster imagery observers expect the vector output to be an accurate representation of the raster input. However, perception studies suggest that viewers implicitly account for the lossy nature of the rasterization process and mentally smooth and simplify the observed boundaries. Our core algorithmic challenge is to balance these conflicting cues and obtain a piecewise continuous vectorization whose discontinuities, or corners, are aligned with human expectations.

Authors' addresses: Shayan Hoshayari, University of British Columbia, hoshayari@cs.ubc.ca; Edoardo Alberto Dominici, University of British Columbia, dedoardo@cs.ubc.ca; Alla Sheffer, University of British Columbia, sheffa@cs.ubc.ca; Nathan Carr, Adobe, ncarr@adobe.com; Zhaowen Wang, Adobe, zhawang@adobe.com; Duygu Ceylan, Adobe, duygu.ceylan@gmail.com; I-Chao Shen, National Taiwan University, jdiylshen@gmail.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART118 \$15.00

<https://doi.org/10.1145/3197517.3201312>

Our framework centers around a simultaneous spline fitting and corner detection method that combines a learned metric, that approximates human perception of boundary discontinuities on raster inputs, with perception-driven algorithmic discontinuity analysis. The resulting method balances local cues provided by the learned metric with global cues obtained by balancing simplicity and continuity expectations. Given the finalized set of corners, our framework connects those using simple, continuous curves that capture input regularities. We demonstrate our method on a range of inputs and validate its superiority over existing alternatives via an extensive comparative user study.

CCS Concepts: • **Computing methodologies** → **Image manipulation**;

Additional Key Words and Phrases: clip-art, vectorization

## ACM Reference Format:

Shayan Hoshayari, Edoardo Alberto Dominici, Alla Sheffer, Nathan Carr, Zhaowen Wang, Duygu Ceylan, and I-Chao Shen. 2018. Perception-Driven Semi-Structured Boundary Vectorization. *ACM Trans. Graph.* 37, 4, Article 118 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201312>

## 1 INTRODUCTION

Many artist-generated images frequently used in digital artwork (such as computer icons, comic book imagery, and simple graphic illustrations) consist of distinctly colored regions with piecewise continuous boundaries and visually pronounced corners (Figure 1). Such *semi-structured* images can be compactly represented in vector form enabling editing and other manipulation. For a range of historic reasons, many such images are still generated in raster, or bitmap formats; large legacy collections of semi-structured images and artwork exist in raster formats only, creating a demand for robust vectorization algorithms. For instance, the Adobe Stock Image database contains over nine million raster images labeled as icons or clip-art. Converting these images into a vector format

# Fast and Deep Deformation Approximations

STEPHEN W. BAILEY, University of California, Berkeley

DAVE OTTE, DreamWorks Animation

PAUL DILORENZO, DreamWorks Animation

JAMES F. O'BRIEN, University of California, Berkeley

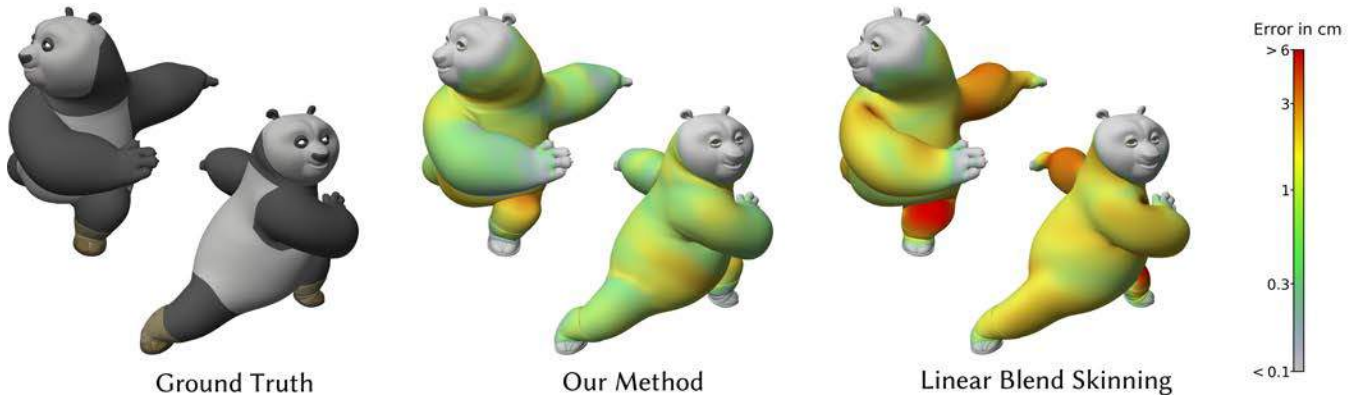


Fig. 1. Comparison of a deformed mesh using a fully evaluated rig, our fast deformation approximation, and linear blend skinning. The meshes are colored to indicate the distance error for each vertex compared with the ground truth mesh.

Character rigs are procedural systems that compute the shape of an animated character for a given pose. They can be highly complex and must account for bulges, wrinkles, and other aspects of a character's appearance. When comparing film-quality character rigs with those designed for real-time applications, there is typically a substantial and readily apparent difference in the quality of the mesh deformations. Real-time rigs are limited by a computational budget and often trade realism for performance. Rigs for film do not have this same limitation, and character riggers can make the rig as complicated as necessary to achieve realistic deformations. However, increasing the rig complexity slows rig evaluation, and the animators working with it can become less efficient and may experience frustration. In this paper, we present a method to reduce the time required to compute mesh deformations for film-quality rigs, allowing better interactivity during animation authoring and use in real-time games and applications. Our approach learns the deformations from an existing rig by splitting the mesh deformation into linear and nonlinear portions. The linear deformations are computed directly from the transformations of the rig's underlying skeleton. We use deep learning methods to approximate the remaining nonlinear portion. In the examples we show from production rigs used to animate lead characters, our approach reduces the computational time spent on evaluating deformations by a factor of 5 $\times$ -10 $\times$ . This significant savings allows us to

run the complex, film-quality rigs in real-time even when using a CPU-only implementation on a mobile device.

CCS Concepts: • **Computing methodologies** → **Neural networks; Animation**;

Additional Key Words and Phrases: character rigs, deep learning, mesh deformations

## ACM Reference Format:

Stephen W. Bailey, Dave Otte, Paul DiLorenzo, and James F. O'Brien. 2018. Fast and Deep Deformation Approximations. *ACM Trans. Graph.* 37, 4, Article 119 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201300>

## 1 INTRODUCTION

The level of detail that can be included in character rigs for interactive applications such as video games and virtual reality is limited by computational costs. These types of rigs need to run at interactive rates, and therefore need to be evaluated as quickly as possible. Because of this limitation, real-time character rigs often lack a high level of detail. In contrast, film-quality character rigs are not limited by hard computational constraints and their mesh deformations appear more detailed and complex. Unfortunately, film-quality character rigs are not suitable for real-time applications. A single film-quality rig might be able to run at interactive rates on a high-end machine, but typically only after tremendous effort has been spent to optimize the rig, parallelize its evaluation, and shift parts of the computation to the high-end machine's GPU. We would like to use these high quality rigs to increase the detail of characters in interactive applications, particularly those running on modest computing platforms such as mobile devices or game consoles. However, directly plugging these computationally intensive rigs into an interactive application is generally infeasible. The

Authors' addresses: Stephen W. Bailey, University of California, Berkeley; Dave Otte, DreamWorks Animation; Paul DiLorenzo, DreamWorks Animation; James F. O'Brien, University of California, Berkeley.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART119 \$15.00 <https://doi.org/10.1145/3197517.3201300>



# Numerical Coarsening using Discontinuous Shape Functions

JIONG CHEN, State Key Lab of CAD&CG, Zhejiang University  
HUJUN BAO\*, State Key Lab of CAD&CG, Zhejiang University  
TIANYU WANG, State Key Lab of CAD&CG, Zhejiang University  
MATHIEU DESBRUN, Caltech  
JIN HUANG\*, State Key Lab of CAD&CG, Zhejiang University

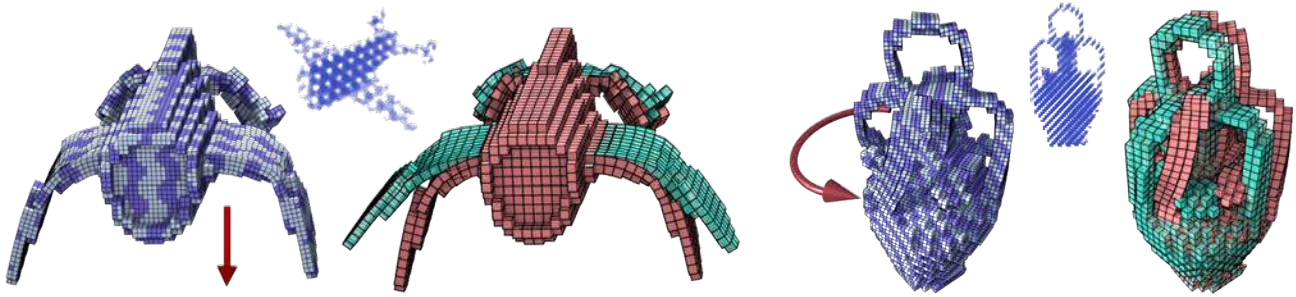


Fig. 1. **Numerical coarsening of nonlinear materials.** We introduce a new approach for numerically capturing the behavior of inhomogeneous and nonlinear materials on coarse grids. For two composite models (left: plane with downwards forces exerted on wings; right: vase being twisted) made of two different compressible neo-Hookean materials discretized on a fine mesh (Poisson ratio: 0.45; Young's moduli:  $1e3$  and  $5e4$ ; volume rendering insets illustrate the spatial distribution of their composite material), our coarsened model with eight times less elements (red) perfectly reproduces the original behavior, while a regular FEM simulation using trilinear shape functions (green) on the same coarse mesh leads to drastic overstiffening.

In this paper, an efficient and scalable approach for simulating inhomogeneous and non-linear elastic objects is introduced. Our numerical coarsening approach consists in optimizing non-conforming and matrix-valued shape functions to allow for predictive simulation of heterogeneous materials with non-linear constitutive laws even on coarse grids, thus saving orders of magnitude in computational time compared to traditional finite element computations. The set of local shape functions over coarse elements is carefully tailored in a preprocessing step to balance geometric continuity and local material stiffness. In particular, we do not impose continuity of our material-aware shape functions between neighboring elements to significantly reduce the fictitious numerical stiffness that conforming bases induce; however, we enforce crucial geometric and physical properties such as partition of unity and exact reproduction of representative fine displacements to eschew the use of discontinuous Galerkin methods. We demonstrate that we can simulate, with no parameter tuning, inhomogeneous and non-linear materials significantly better than previous approaches that traditionally try to homogenize the constitutive model instead.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: Numerical coarsening, nonlinear constitutive models, deformable body simulation.

\* Corresponding authors: {bao, hj}@cad.zju.edu.cn.

Authors' addresses: J. Chen, H. Bao, T. Wang, J. Huang, Zhejiang University, China; M. Desbrun, California Institute of Technology, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART120 \$15.00

<https://doi.org/10.1145/3197517.3201386>

## ACM Reference Format:

Jiong Chen, Hujun Bao\*, Tianyu Wang, Mathieu Desbrun, and Jin Huang\*. 2018. Numerical Coarsening using Discontinuous Shape Functions. *ACM Trans. Graph.* 37, 4, Article 120 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201386>

## 1 INTRODUCTION

Efficacy in simulating complex deformable models is a long standing goal in computer animation. As the demand for ever larger and more complex simulation increases, so does the need for more elaborate numerical methods whose computational costs grows only slowly with the structural complexity of the object being simulated. Traditional animation methods can only capture the proper behavior of a physical object if one uses a mesh fine enough to resolve the small-scale heterogeneities, leading to prohibitive simulation time when the fine scales are geometrically complex or have a drastically different stiffness (such as veinal structures in an organ or microstructures in a metamaterial). Simply ignoring these fine scales dramatically affects the overall dynamics of the object: a coarse simulation can fail to capture even the simplest deformation, at times rendering the object dramatically more rigid than in reality—an archetypical example being the issue of locking in finite element modeling, where incompressibility is at odds with the conforming constraints of low-order finite elements. As visual and manufacturing fidelity calls for an increasingly widespread use of non-linear geometric models and non-linear constitutive laws, the design of numerical algorithms that combine efficiency and scalability is increasingly pressing in graphics and computational physics.

Over the years, different strategies have been devised to address this enduring problem in simulation. A first family of efforts uses adaptive solvers to counteract the spurious numerical rigidity and

# Magnetization Dynamics for Magnetic Object Interactions

SEUNG-WOOK KIM, Korea University

SUN YOUNG PARK, Korea University

JUNGHYUN HAN, Korea University

The goal of this paper is to simulate the interactions between magnetic objects in a physically correct way. The simulation scheme is based on magnetization dynamics, which describes the temporal change of magnetic moments. For magnetization dynamics, the Landau-Lifshitz-Gilbert equation is adopted, which is widely used in micromagnetics. Through effectively-designed novel models of magnets, it is extended into the macro scale so as to be combined with real-time rigid-body dynamics. The overall simulation is stable and enables us to implement mutual induction and remanence that have not been tackled by the state-of-the-art technique in magnet simulation. The proposed method can be applied to various fields including magnet experiments in the virtual world.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: Magnetic fields and forces, magnetization, rigid body dynamics

## ACM Reference Format:

Seung-wook Kim, Sun Young Park, and JungHyun Han. 2018. Magnetization Dynamics for Magnetic Object Interactions. *ACM Trans. Graph.* 37, 4, Article 121 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201402>

## 1 INTRODUCTION

The last three decades have seen the prosperity of physics-based simulation in computer graphics. The results have been used in many areas including video games, visual effects in films, and education. A unique subfield of the physics-based simulation is focused on magnets. In the pioneering work of [Thomaszewski et al. 2008], the force and torque between magnetic objects were computed using the Maxwell's equations and integrated into rigid-body dynamics. Thanks to their efforts, magnetic simulation has recently begun to be adopted in education and art, as illustrated in Fig. 1. However, the potential demand of magnetic simulation requires the current techniques to advance one step further and address a few critical issues:

- **Magnetic mutual induction:** The magnetic induction of an object is determined by *all* magnetic objects in the environment. Fig. 2-(a) shows an example captured in a real-world experiment. Not only the bar magnet but also the iron ball in direct contact with it induces the second ball. The other balls are similarly induced by the balls preceding them and the bar magnet. Thus, we

Authors' addresses: Seung-wook Kim, Korea University, [wook0249@korea.ac.kr](mailto:wook0249@korea.ac.kr); Sun Young Park, Korea University, [iamsun0720@korea.ac.kr](mailto:iamsun0720@korea.ac.kr); JungHyun Han, Korea University, [jhan@korea.ac.kr](mailto:jhan@korea.ac.kr).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART121 \$15.00

<https://doi.org/10.1145/3197517.3201402>

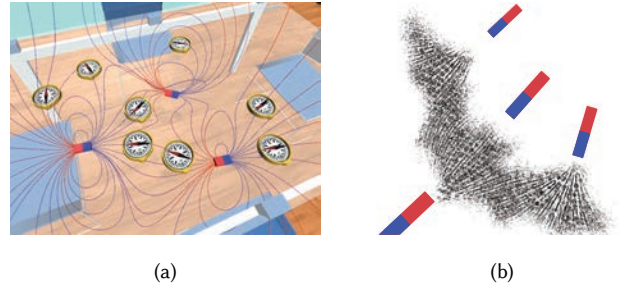


Fig. 1. The state of the art in magnet simulation: (a) Magnetic field lines and compasses [Park et al. 2016]. (b) Iron filing art [Yoon et al. 2014].

have a chain of induced magnets. Unfortunately, such mutual induction has not been tackled by the state-of-the-art techniques.

- **Magnetic remanence:** Many induced objects have a tendency of preserving magnetization for a certain period of time. Fig. 2-(b) shows a well-known example of first magnetizing a metallic material and then using its remanence to magnetize other objects. There has been no attempt to implement this phenomenon.
- **Simulation stability:** A magnetic object exposed to a magnetic field is accelerated by a magnetic force. In general, the closer the magnetic objects are to each other, the more induced they become, and thus the magnetic forces that they produce become stronger. This often causes the magnetic objects to interpenetrate each other and produce much stronger magnetic forces, which eventually make the simulation unstable. In order to resolve this problem, a smaller time step size can be used in the rigid-body simulation. However, this is not appropriate for real-time simulation.

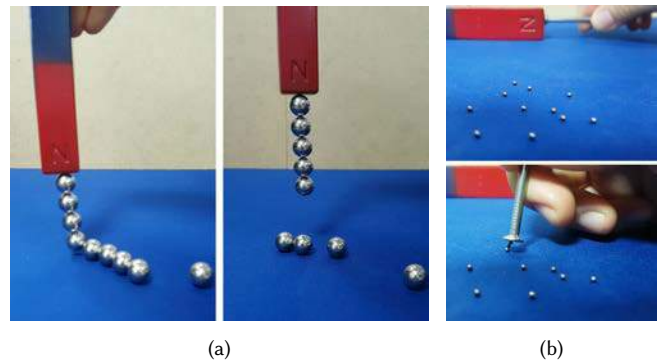


Fig. 2. Indispensable components of virtual magnet experiments: (a) Mutual induction. (b) Remanence.

# Visual Rhythm and Beat

ABE DAVIS, Stanford University

MANEESH AGRAWALA, Stanford University

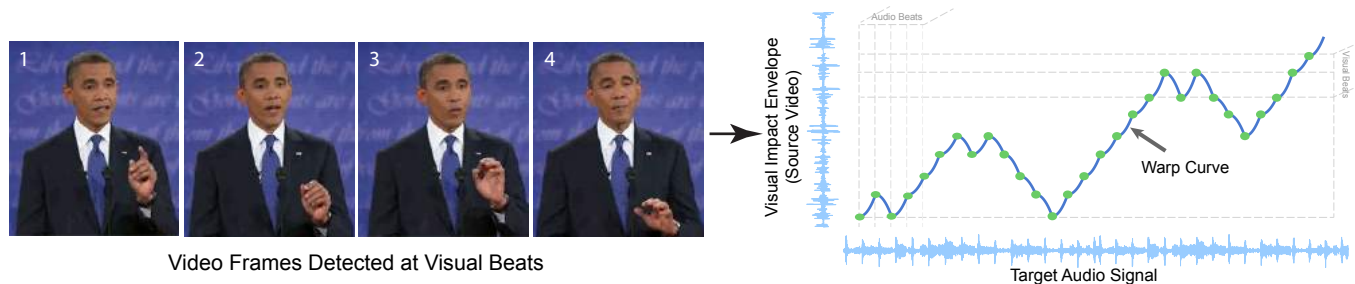


Fig. 1. **Accidental Dance and Unfolding** - we analyze visual rhythm in a collection of video to search for segments of unintentionally-rhythmic motion. Then, we warp those segments into precise temporal alignment with a musical target to synthesize dance performances. On the left we see video frames corresponding to the moments of four consecutive visual beats detected in a 2012 presidential debate video [WSJDigitalNetwork 2012]. These visual beats lie at the high and low points of a repetitive up-and-down hand gesture. On the right is a warp curve showing the process of unfolding, which synthesizes dance video corresponding to a random walk through the visual beats of a source segment.

We present a visual analogue for musical rhythm derived from an analysis of motion in video, and show that alignment of visual rhythm with its musical counterpart results in the appearance of dance. Central to our work is the concept of visual beats — patterns of motion that can be shifted in time to control visual rhythm. By warping visual beats into alignment with musical beats, we can create or manipulate the appearance of dance in video. Using this approach we demonstrate a variety of retargeting applications that control musical synchronization of audio and video: we can change what song performers are dancing to, warp irregular motion into alignment with music so that it appears to be dancing, or search collections of video for moments of accidentally dance-like motion that can be used to synthesize musical performances.

CCS Concepts: • **Computing methodologies** → *Computational photography; Image processing; Image-based rendering*; • **Human-centered computing** → Graphics input devices; Sound-based input / output;

Additional Key Words and Phrases: music, dance, video editing

## ACM Reference Format:

Abe Davis and Maneesh Agrawala. 2018. Visual Rhythm and Beat. *ACM Trans. Graph.* 37, 4, Article 122 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201371>

Authors' addresses: Abe Davis, Stanford University, [abedavis@stanford.edu](mailto:abedavis@stanford.edu); Maneesh Agrawala, Stanford University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART122 \$15.00

<https://doi.org/10.1145/3197517.3201371>

## 1 INTRODUCTION

Music and dance are closely related through the concept of rhythm, which describes how events—e.g., the sound of an instrument or the movement of a body—are distributed in time. Rhythm is in some sense a very intuitive concept: infants can recognize and follow basic rhythms at as early as six months of age [Cirelli et al. 2016; Repp and Su 2013], and even some animals—certain parrots and elephants, for example—are known to move in time with simple music [Patel and Demorest 2013; Patel et al. [n. d.]]. However, the task of quantifying rhythm is not trivial, and has been the topic of extensive research in the context of both music [Böck and Gerhard Widmer 2013; Dixon 2006; Ellis 2007; Goto 2002; Grosche et al. 2010; Hu et al. 2017; Lerch 2012] and dance [Brick and Boker 2011; Dyaberi et al. 2006]. Our work builds on that research to explore a visual analogue for rhythm—which we call **visual rhythm**—in video. Just as musical rhythm captures the temporal arrangement of sounds, visual rhythm captures the temporal arrangement of visible motion. We focus on analyzing that motion to identify structure related to dance.

Our central hypothesis is that music and dance are characterized by complementary rhythmic structure in audible and visible signals. Our exploration of that structure builds on the concept of **visual beats**—visual events that, when temporally aligned with musical beats, create the appearance of dance. The relationship between visual and musical beats provides a starting point from which we derive visual analogues for other rhythmic concepts, including onset strength and tempo. Visual beats also give us a recipe for manipulating rhythmic structure in video: we first identify visual beats, then time-warp those beats into alignment with a specified target. Provided we are able to identify the necessary beats, we show that it is possible to warp video into dance-like alignment with any song of our choice.



# Perception-Aware Modeling and Fabrication of Digital Drawing Tools

MICHAL PIOVARČI, MPI Informatik, Saarland University, MMCI, and Università della Svizzera italiana, Switzerland

DAVID I. W. LEVIN, University of Toronto, Canada

DANNY M. KAUFMAN, Adobe Research, United States of America

PIOTR DIDYK, MPI Informatik, Saarland University, MMCI, and Università della Svizzera italiana, Switzerland

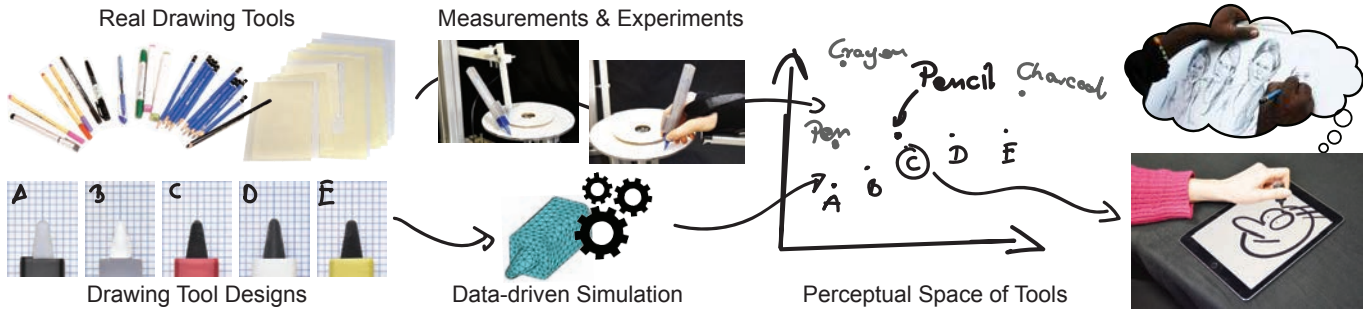


Fig. 1. We propose a system for fabricating digital drawing tools that mimic the feel of real tools. To this end, we measure properties of different real drawing tools, study their perception, and design a perception-aware space of drawing tools. We later develop a simulation technique which allows us to embed new designs into the space, evaluate the pairwise similarity between them and the tools we want to replicate. This drives the design process of different digital tools.

Digital drawing is becoming a favorite technique for many artists. It allows for quick swaps between different materials, reverting changes, and applying selective modifications to finished artwork. These features enable artists to be more efficient and creative. A significant disadvantage of digital drawing is poor haptic feedback. Artists are usually limited to one surface and a few different stylus nibs, and while they try to find a combination that suits their needs, this is typically challenging. In this work, we address this problem and propose a method for designing, evaluating, and optimizing different stylus designs. We begin with collecting a representative set of traditional drawing tools. We measure their physical properties and conduct a user experiment to build a perceptual space that encodes perceptually-relevant attributes of drawing materials. The space is optimized to both explain our experimental data and correlate it with measurable physical properties. To embed new drawing tool designs into the space without conducting additional experiments and measurements, we propose a new, data-driven simulation technique for characterizing stylus-surface interaction. We finally leverage the perceptual space, our simulation, and recent advancements in multi-material 3D printing to demonstrate the application of our system in the design of new digital drawing tools that mimic traditional drawing materials.

Authors' addresses: Michal Piovarči, MPI Informatik, Saarland University, MMCI, Università della Svizzera italiana, Switzerland; David I. W. Levin, University of Toronto, Canada; Danny M. Kaufman, Adobe Research, United States of America; Piotr Didyk, MPI Informatik, Saarland University, MMCI, Università della Svizzera italiana, Switzerland.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART123 \$15.00

<https://doi.org/10.1145/3197517.3201322>

CCS Concepts: • **Computing methodologies** → **Graphics input devices; Perception; Physical simulation**; • **Human-centered computing** → **Haptic devices**;

Additional Key Words and Phrases: digital drawing, haptic, perception, stylus, contact simulation

## ACM Reference Format:

Michal Piovarči, David I. W. Levin, Danny M. Kaufman, and Piotr Didyk. 2018. Perception-Aware Modeling and Fabrication of Digital Drawing Tools. *ACM Trans. Graph.* 37, 4, Article 123 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201322>

## 1 INTRODUCTION

Everyone draws and everyone writes: some with skill and precision, while many of us simply aim to communicate and record. All of us begin by writing and drawing with simple yet highly effective hand tools, e.g. crayons, pencils and pens, technologies that have largely remained unchanged for well over a century [Petroski 1992; Van Dulken 2002]. As we turn to digital devices, styli and tablets offer the promise of integrating our ease and comfort in the affordances of hand implements with the advantages of software tools.

Many previous obstacles to stylus adoption are gone. Stylus response time, pressure sensitivity, stroke capture and palm rejection have all been rapidly improving. Concurrently, drawing apps like *Paper by FiftyThree*<sup>1</sup> have successfully focused on replicating the wide variety of physical marks made by differing implements and materials. However, the complementary goal of emulating the physical *feel* of real-world drawing and writing materials with styli remains an important and ongoing challenge.

<sup>1</sup><https://www.fiftythree.com/>

# Denoising with Kernel Prediction and Asymmetric Loss Functions

THIJS VOGELS, Disney Research

FABRICE ROUSSELLE, Disney Research

BRIAN MCWILLIAMS, Disney Research

GERHARD RÖTHLIN, Disney Research

ALEX HARVILL, Pixar Animation Studios

DAVID ADLER, Walt Disney Animation Studios

MARK MEYER, Pixar Animation Studios

JAN NOVÁK, Disney Research



Fig. 1. We introduce a novel, modular architecture based on kernel-predicting neural networks that performs multi-scale, temporal denoising of rendered sequences. Our network can be retrained using a small amount of data and performs robustly with artist-control over the variance-bias tradeoff on novel data, such as these two examples from our test set. © Disney / Pixar, © Disney.

We present a modular convolutional architecture for denoising rendered images. We expand on the capabilities of kernel-predicting networks by combining them with a number of task-specific modules, and optimizing the assembly using an asymmetric loss. The source-aware encoder—the first module in the assembly—extracts low-level features and embeds them into a common feature space, enabling quick adaptation of a trained network to novel data. The spatial and temporal modules extract abstract, high-level features for kernel-based reconstruction, which is performed at three different spatial scales to reduce low-frequency artifacts. The complete network is trained using a class of asymmetric loss functions that are designed to preserve details and provide the user with a direct control over the variance-bias trade-off during inference. We also propose an error-predicting module for inferring reconstruction error maps that can be used to drive adaptive sampling. Finally, we present a theoretical analysis of convergence rates of kernel-predicting architectures, shedding light on why kernel prediction performs better than synthesizing the colors directly, complementing the

Authors' addresses: Disney Research, Stampfenbachstrasse 48, Zürich, Switzerland; Pixar Animation Studios, 1200 Park Ave, Emeryville, CA, USA; Walt Disney Animation Studios, 2100 W. Riverside Drive, Burbank, CA, USA. Corresponding author: Jan Novák, Disney Research, jan.novak@disneyresearch.com.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART124

<https://doi.org/10.1145/3197517.3201388>

empirical evidence presented in this and previous works. We demonstrate that our networks attain results that compare favorably to state-of-the-art methods in terms of detail preservation, low-frequency noise removal, and temporal stability on a variety of production and academic datasets.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; *Rendering*; Ray tracing;

Additional Key Words and Phrases: Monte Carlo denoising, neural networks, kernel prediction, asymmetric loss, temporal, multi-scale

## ACM Reference Format:

Thijs Vogels, Fabrice Rousselle, Brian McWilliams, Gerhard Röthlin, Alex Harvill, David Adler, Mark Meyer, and Jan Novák. 2018. Denoising with Kernel Prediction and Asymmetric Loss Functions. *ACM Trans. Graph.* 37, 4, Article 124 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201388>

## 1 INTRODUCTION

Monte Carlo (MC) rendering is used ubiquitously in computer animation and visual effect productions [Keller et al. 2015]. Despite continuously increasing computational power, the cost of constructing light paths—the core component of image synthesis—remains a limiting practical constraint that leads to noise. Among the many strategies explored to reduce Monte Carlo noise, image-space denoising has emerged as a particularly attractive solution due to its effectiveness and ease of integration into rendering pipelines.

Until recently, the best-performing MC denoisers were hand-designed and based on linear regression models [Zwicker et al.

# Bayesian online regression for adaptive direct illumination sampling

PETR VÉVODA\*, Charles University, Prague and Render Legion, a. s.

IVO KONDAPANENI\*, Charles University, Prague

JAROSLAV KŘIVÁNEK, Charles University, Prague and Render Legion, a. s.

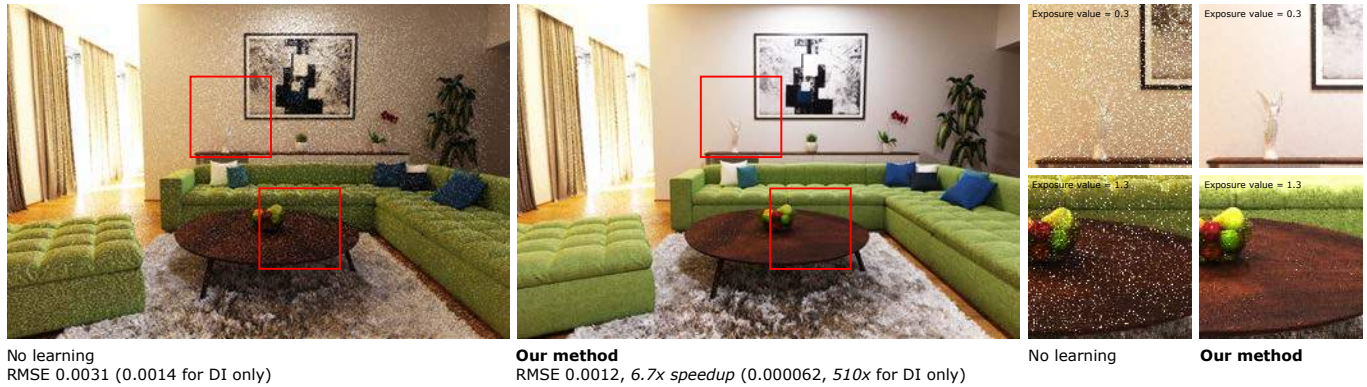


Fig. 1. Equal-time comparison (60 s) of path-traced global illumination solutions computed using our learning-based direct illumination sampling method (right) and a baseline sampling method without learning (left). While both methods start off by sampling lights proportionally to rough estimates of their unoccluded contribution, our method progressively incorporates information about their actual contributions, including visibility, dramatically reducing image variance.

Direct illumination calculation is an important component of any physically-based renderer with a substantial impact on the overall performance. We present a novel adaptive solution for unbiased Monte Carlo direct illumination sampling, based on online learning of the light selection probability distributions. Our main contribution is a formulation of the learning process as Bayesian regression, based on a new, specifically designed statistical model of direct illumination. The net result is a set of regularization strategies to prevent over-fitting and ensure robustness even in early stages of calculation, when the observed information is sparse. The regression model captures spatial variation of illumination, which enables aggregating statistics over relatively large scene regions and, in turn, ensures a fast learning rate. We make the method scalable by adopting a light clustering strategy from the Lightcuts method, and further reduce variance through the use of control variates. As a main design feature, the resulting algorithm is virtually free of any preprocessing, which enables its use for interactive progressive rendering, while the online learning still enables super-linear convergence.

CCS Concepts: • **Computing methodologies** → **Rendering**; *Visibility*; *Machine learning*;

Authors' addresses: Petr Vévoda, Charles University, Prague, Malostranské náměstí 25, Prague, 11800, Render Legion, a. s. Karlovo náměstí 288/17, Prague, 12000, petrvevoda@seznam.cz; Ivo Kondapaneni, Charles University, Prague, Malostranské náměstí 25, Prague, 11800, ivo.kondapaneni@gmail.com; Jaroslav Krivánek, Charles University, Prague, Malostranské náměstí 25, Prague, 11800, Render Legion, a. s. Karlovo náměstí 288/17, Prague, 12000, jaroslav.krivanek@mff.cuni.cz.

\*Petr Vévoda and Ivo Kondapaneni share the first authorship of this work. Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART125 \$15.00

<https://doi.org/10.1145/3197517.3201340>

Additional Key Words and Phrases: direct illumination, adaptive sampling, visibility, learning.

## ACM Reference Format:

Petr Vévoda, Ivo Kondapaneni, and Jaroslav Krivánek. 2018. Bayesian online regression for adaptive direct illumination sampling. *ACM Trans. Graph.* 37, 4, Article 125 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201340>

## 1 INTRODUCTION

Realistic rendering today is almost entirely based on Monte Carlo (MC) methods. The *indirect* illumination component has traditionally been held responsible for the undesirable image noise produced by such algorithms, which is probably why the *direct* illumination has received disproportionately less attention in research. However, many scenes in digital production feature complex lighting setups, and practical experience shows that it is often direct illumination that is responsible for the majority of image noise.

In this paper, we aim at unbiased direct illumination estimation for MC renderers. Specifically, we address the problem of randomly choosing an appropriate light source for a given scene location, so that variance of the direct illumination estimator is minimized. This could be achieved by choosing lights with probability proportional to their respective contributions, but these are unknown at the outset, they are costly to evaluate and difficult to predict. This is true especially due to the visibility, since it can be discontinuous and its evaluation involves expensive ray casting.

One possible solution would involve constructing the light sampling distributions in a preprocessing step [Georgiev et al. 2012]. However, long preprocessing disqualifies any form of *interactive rendering* – a crucial feature of any modern progressive renderer, a feature that we consider a hard constraint in our work. Such preprocessing can be avoided by learning from the observed samples



# Deep Image-Based Relighting from Optimal Sparse Samples

ZEXIANG XU, University of California, San Diego

KALYAN SUNKAVALLI, Adobe Research

SUNIL HADAP, Adobe Research

RAVI RAMAMOORTHY, University of California, San Diego

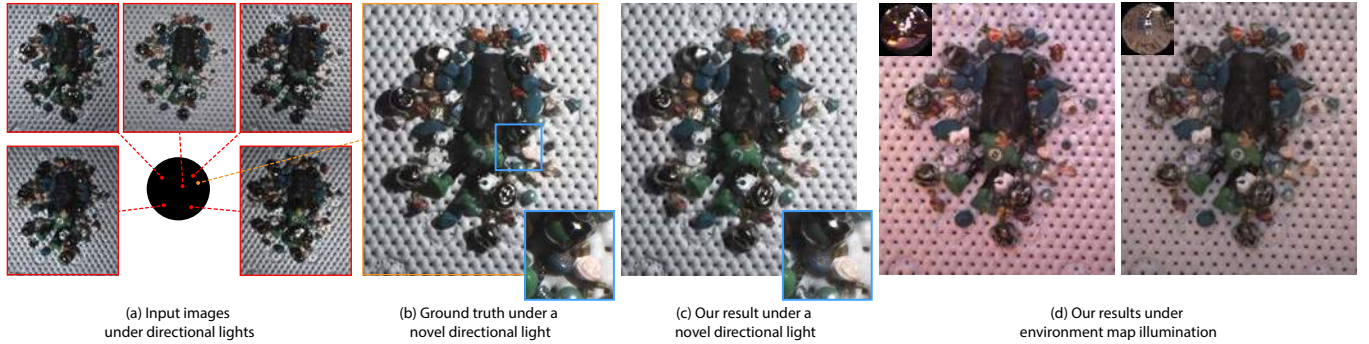


Fig. 1. We propose a learning-based method that takes only five images of a scene under directional lights (a, light directions marked on circle in red) and reconstructs its appearance (c) under a novel directional light in the upper hemisphere (marked in orange). Our method trains a fully-convolutional neural network to jointly learn the optimal input light directions and relighting function for any scene. The network can reconstruct even high-frequency patterns like specular shading and cast shadows (insets in c) and produces photorealistic relighting results that closely match the ground truth (b). Moreover, by generating images for every direction in the upper hemisphere, our method can be used to relight scenes under environment map illumination (d).

We present an image-based relighting method that can synthesize scene appearance under novel, distant illumination from the visible hemisphere, from only five images captured under pre-defined directional lights. Our method uses a deep convolutional neural network to regress the relit image from these five images; this *relighting network* is trained on a large synthetic dataset comprised of procedurally generated shapes with real-world reflectances. We show that by combining a custom-designed *sampling network* with the relighting network, we can jointly learn both the optimal input light directions and the relighting function. We present an extensive evaluation of our network, including an empirical analysis of reconstruction quality, optimal lighting configurations for different scenarios, and alternative network architectures. We demonstrate, on both synthetic and real scenes, that our method is able to reproduce complex, high-frequency lighting effects like specularities and cast shadows, and outperforms other image-based relighting methods that require an order of magnitude more images.

CCS Concepts: • **Computing methodologies** → **Rendering**;

Additional Key Words and Phrases: Image-based relighting, Illumination, Convolutional Neural Network, Sparse sampling, Appearance capture

Authors' addresses: Zexiang Xu, University of California, San Diego, zexiangxu@cs.ucsd.edu; Kalyan Sunkavalli, Adobe Research, sunkaval@adobe.com; Sunil hadap, Adobe Research, hadap@adobe.com; Ravi Ramamoorthi, University of California, San Diego, ravir@cs.ucsd.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://doi.org/10.1145/3197517.3201313).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART126 \$15.00

<https://doi.org/10.1145/3197517.3201313>

## ACM Reference Format:

Zexiang Xu, Kalyan Sunkavalli, Sunil hadap, and Ravi Ramamoorthi. 2018. Deep Image-Based Relighting from Optimal Sparse Samples. *ACM Trans. Graph.* 37, 4, Article 126 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201313>

## 1 INTRODUCTION

Rendering a scene under novel lighting is a long-studied vision and graphics problem with applications in visual effects, virtual and augmented reality, and product visualization for e-commerce. One approach to relighting is to reconstruct the geometry and material properties of the scene, and render this reconstruction under novel lighting. However, reconstruction is an extremely challenging problem, especially for scenes with complex geometry and reflectance.

Image-based relighting methods bypass reconstruction by directly modeling the scene's light transport function. Assuming distant illumination, the light transport function,  $T(\mathbf{x}, \omega)$ , maps incident illumination from direction  $\omega$  to outgoing radiance at pixel  $x$  (towards the camera), and allows for the scene to be rendered under novel distant lighting as:

$$I(\mathbf{x}) = \int_{\mathcal{L}} T(\mathbf{x}, \omega) L(\omega) d\omega, \quad (1)$$

where  $L(\omega)$  is the radiance of the incident illumination from direction  $\omega$ . The light transport function can be sampled by capturing images under different lighting conditions; for example, an image of the scene under a single directional light from direction  $\omega_j$ , yields the sample:  $I_j(\cdot) = T(\cdot, \omega_j)$ . Image-based relighting methods use a set of such samples,  $\{I_j, \omega_j \mid j = 1, 2, \dots, k\}$ , to reproduce scene appearance,  $I_n$ , under a novel light,  $\omega_n$ . Because the light transport

# Efficient Reflectance Capture Using an Autoencoder

KAIZHANG KANG, State Key Lab of CAD&CG, Zhejiang University, China

ZIMIN CHEN, State Key Lab of CAD&CG, Zhejiang University, China

JIAPING WANG, Sinovation Ventures, USA

KUN ZHOU, State Key Lab of CAD&CG, Zhejiang University, China

HONGZHI WU, State Key Lab of CAD&CG, Zhejiang University, China

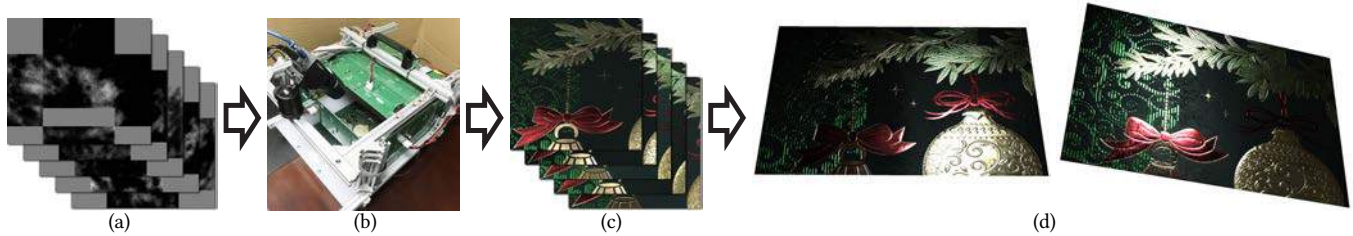


Fig. 1. Using as few as 16 ~ 32 lighting patterns (a) learned from a large amount of reflectance data, we efficiently capture photographs (c) of a planar physical sample in a mini, near-field lightstage (b), and faithfully reconstruct its SVBRDF that can be rendered under novel lighting and view conditions (d).

We propose a novel framework that automatically learns the lighting patterns for efficient reflectance acquisition, as well as how to faithfully reconstruct spatially varying anisotropic BRDFs and local frames from measurements under such patterns. The core of our framework is an asymmetric deep autoencoder, consisting of a nonnegative, linear encoder which directly corresponds to the lighting patterns used in physical acquisition, and a stacked, nonlinear decoder which computationally recovers the BRDF information from captured photographs. The autoencoder is trained with a large amount of synthetic reflectance data, and can adapt to various factors, including the geometry of the setup and the properties of appearance. We demonstrate the effectiveness of our framework on a wide range of physical materials, using as few as 16 ~ 32 lighting patterns, which correspond to 12 ~ 25 seconds of acquisition time. We also validate our results with the ground truth data and captured photographs. Our framework is useful for increasing the efficiency in both novel and existing acquisition setups.

CCS Concepts: • **Computing methodologies** → **Reflectance modeling**;

Additional Key Words and Phrases: reflectance acquisition, optimal sampling, lighting patterns, SV-BRDF

Hongzhi Wu (hwu@acm.org) is the first corresponding author, and Kun Zhou (kunzhou@acm.org) is the second. Kaizhang Kang and Zimin Chen contributed equally. Authors' addresses: Kaizhang Kang, State Key Lab of CAD&CG, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, China; Zimin Chen, State Key Lab of CAD&CG, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, China; Jiaping Wang, Sinovation Ventures, 4158 Interdale Wy, Palo Alto, CA, 94306, USA; Kun Zhou, State Key Lab of CAD&CG, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, China; Hongzhi Wu, State Key Lab of CAD&CG, Zhejiang University, 866 Yuhangtang Rd, Hangzhou, 310058, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART127 \$15.00

<https://doi.org/10.1145/3197517.3201279>

## ACM Reference Format:

Kaizhang Kang, Zimin Chen, Jiaping Wang, Kun Zhou, and Hongzhi Wu. 2018. Efficient Reflectance Capture Using an Autoencoder. *ACM Trans. Graph.* 37, 4, Article 127 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201279>

## 1 INTRODUCTION

Digitally acquiring high-quality material appearance from the real-world is a crucial and challenging problem, with applications in visual effects, e-commerce, product design and entertainment. One fundamental difficulty here is the sheer complexity of the material appearance. It can be modeled as a Spatially Varying Bidirectional Reflectance Distribution Function (SVBRDF), a 6D function that varies with location, lighting and view directions. Directly capturing the unknown, general SVBRDF of a planar sample with a conventional gonioreflectometer [Dana et al. 1999; Lawrence et al. 2006] requires taking thousands or even millions of photographs from all possible combinations of lighting and view directions, in order to sufficiently sample high-frequency features such as sharp highlights. This makes the approach prohibitively expensive both in time and storage.

Significant research efforts have been made to improve the efficiency of reflectance acquisition over the past years. One popular class of methods for high-quality SVBRDF acquisition are based on complex lighting patterns. Instead of using a single light source at a time, such a system illuminates a planar physical sample with many intensity-programmable light sources simultaneously, effectively forming different lighting patterns. It then takes corresponding photographs with a camera, and reconstructs the reflectance from these measurements. Representative work includes the lightstage [Ghosh et al. 2009; Tunwattanapong et al. 2013], the linear light source reflectometry [Chen et al. 2014; Gardner et al. 2003], and the LCD-based setup [Aittala et al. 2013]. However, hundreds of photographs are still needed in the general setting [Chen et al. 2014], limiting the

# Single-Image SVBRDF Capture with a Rendering-Aware Deep Network

VALENTIN DESCHAI NTRE, Inria, Optis, Université Côte d’Azur

MIIKA AITTALA, MIT CSAIL

FREDO DURAND, MIT CSAIL, Inria, Université Côte d’Azur

GEORGE DRETTAKIS, Inria, Université Côte d’Azur

ADRIEN BOUSSEAU, Inria, Université Côte d’Azur



Fig. 1. From a single flash photograph of a material sample (insets), our deep learning approach predicts a spatially-varying BRDF. See supplemental materials for animations with a moving light.

Texture, highlights, and shading are some of many visual cues that allow humans to perceive material appearance in single pictures. Yet, recovering spatially-varying bi-directional reflectance distribution functions (SVBRDFs) from a single image based on such cues has challenged researchers in computer graphics for decades. We tackle lightweight appearance capture by training a deep neural network to automatically extract and make sense of these visual cues. Once trained, our network is capable of recovering per-pixel normal, diffuse albedo, specular albedo and specular roughness from a single picture of a flat surface lit by a hand-held flash. We achieve this goal by introducing several innovations on training data acquisition and network design. For training, we leverage a large dataset of artist-created, procedural SVBRDFs which we sample and render under multiple lighting directions. We further amplify the data by material mixing to cover a wide diversity of shading effects, which allows our network to work across many material classes. Motivated by the observation that distant regions of a material sample often offer complementary visual cues, we design a network that combines an encoder-decoder convolutional track for local feature extraction with a fully-connected track for *global feature* extraction and propagation. Many important material effects are view-dependent, and as such ambiguous when observed in a single image. We tackle this challenge by defining the loss as a differentiable SVBRDF similarity metric that compares the *renderings* of the predicted maps against renderings of the ground truth from several lighting and viewing directions. Combined together, these novel ingredients bring clear improvement over state of the art methods for single-shot capture of spatially varying BRDFs.

CCS Concepts: • **Computing methodologies** → **Reflectance modeling**; **Image processing**;

Publication rights licensed to ACM. ACM acknowledges that this contribution was authored or co-authored by an employee, contractor or affiliate of a national government. As such, the Government retains a nonexclusive, royalty-free right to publish or reproduce this article, or to allow others to do so, for Government purposes only.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART128 \$15.00

<https://doi.org/10.1145/3197517.3201378>

Additional Key Words and Phrases: material capture, appearance capture, SVBRDF, deep learning

## ACM Reference format:

Valentin Deschaintre, Miika Aittala, Fredo Durand, George Drettakis, and Adrien Bousseau. 2018. Single-Image SVBRDF Capture with a Rendering-Aware Deep Network. *ACM Trans. Graph.* 37, 4, Article 128 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201378>

## 1 INTRODUCTION

The appearance of real-world objects results from complex interactions between light, reflectance, and geometry. Disentangling these interactions is at the heart of *lightweight appearance capture*, which aims at recovering reflectance functions from one or a few photographs of a surface. This task is inherently ill-posed, since many different reflectances can yield the same observed image. For example, any photograph can be perfectly reproduced by a diffuse albedo map, where highlights are “painted” over the surface.

A combination of two strategies is generally employed to deal with this ill-posedness. First, ambiguity can be reduced by collecting additional measurements under different viewing or lighting conditions. While this strategy is currently the most appropriate to achieve high accuracy, it requires precise control of the acquisition process [Xu et al. 2016]. The second strategy is to introduce *a priori* assumptions about the space of plausible solutions. While designing such priors by hand has challenged researchers for decades [Guarnera et al. 2016], Convolutional Neural Networks (CNNs) have emerged as a powerful method to automatically *learn* effective priors from data.

In this paper, we propose a deep learning approach to single-image appearance capture, where we use *forward* rendering simulations to train a neural network to solve the ill-posed *inverse* problem



# Polyhedral Voronoi diagrams for additive manufacturing

JONÀS MARTÍNEZ, SAMUEL HORNUS, HAICHUAN SONG, and SYLVAIN LEFEBVRE, Université de Lorraine, CNRS, Inria, LORIA, France

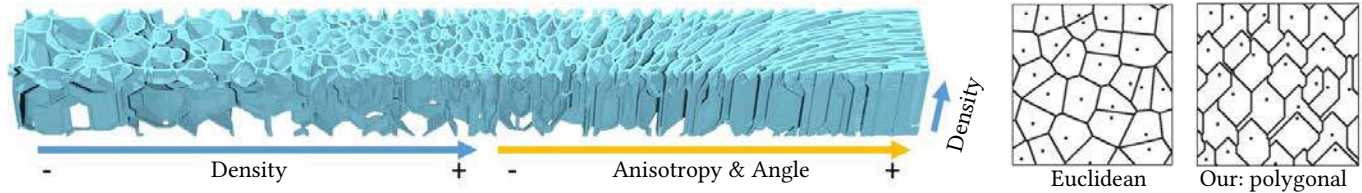


Fig. 1. *Left:* We propose a novel type of microstructures well suited for fabrication on continuous deposition processes such as fused filament fabrication. Our microstructures afford for a wide range of elastic behaviors, from isotropic to orthotropic. As illustrated their geometry can be spatially graded. *Right:* We model the foams as Voronoi diagrams of a special kind, which contrary to Euclidean Voronoi diagrams produce geometry enforcing all fabrication constraints.

A critical advantage of additive manufacturing is its ability to fabricate complex small-scale structures. These microstructures can be understood as a *metamaterial*: they exist at a much smaller scale than the volume they fill, and are collectively responsible for an average elastic behavior different from that of the base printing material making the fabricated object lighter and/or flexible along specific directions. In addition, the average behavior can be graded spatially by progressively modifying the microstructure geometry.

The definition of a microstructure is a careful trade-off between the geometric requirements of manufacturing and the properties one seeks to obtain within a shape: in our case a wide range of elastic behaviors. Most existing microstructures are designed for stereolithography (SLA) and laser sintering (SLS) processes. The requirements are however different than those of continuous deposition systems such as fused filament fabrication (FFF), for which there is currently a lack of microstructures enabling graded elastic behaviors.

In this work we introduce a novel type of microstructures that *strictly enforce* all the requirements of FFF-like processes: continuity, self-support and overhang angles. They offer a range of orthotropic elastic responses that can be graded spatially. This allows to fabricate parts usually reserved to the most advanced technologies on widely available inexpensive printers that also benefit from a continuously expanding range of materials.

CCS Concepts: • **Computing methodologies** → **Shape modeling**;

Additional Key Words and Phrases: Voronoi diagram, 3D printing, additive manufacturing

## ACM Reference Format:

Jonàs Martínez, Samuel Hornus, Haichuan Song, and Sylvain Lefebvre. 2018. Polyhedral Voronoi diagrams for additive manufacturing. *ACM Trans. Graph.* 37, 4, Article 129 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201343>

Authors' address: Jonàs Martínez; Samuel Hornus; Haichuan Song; Sylvain Lefebvre, Université de Lorraine, CNRS, Inria, LORIA, Vandœuvre-les-Nancy, 54600, France 615 Rue du Jardin Botanique.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART129 \$15.00 <https://doi.org/10.1145/3197517.3201343>

## 1 INTRODUCTION

Novel capabilities of additive manufacturing processes are challenging the traditional separation between shape and material. Tool accessibility and mold extraction constraints no longer apply, and shapes with intricate internal details can be fabricated. These internal microstructures modify the large-scale material properties, making parts lighter, porous or flexible and resilient. In addition, these microstructures can be spatially graded: the same object may be rigid where external stresses apply, and lighter in other regions.

This has spawned an intense research effort towards the geometric design of microstructures that can be embedded into volumes while triggering varying properties. The microstructures are typically considered as a *metamaterial*: at large-scale they behave equivalently to a uniform material characterized by a homogeneous elasticity tensor. This average behavior can be analyzed, abstracting away the fine-scale details, in a process called *numerical homogenization*. In addition, the microstructures are often *parametric*: the geometry of their fine scale details is controlled through high-level parameters, such as thickness or orientation. These parameters impact the large-scale elastic behavior, and thus the observed elasticity tensor can be controlled by manipulating these parameters directly. The main interest in using such parametric metamaterials is that their elasticity can be spatially controlled – graded – by a control field, varying mechanical properties within the shape volume. The control field can be either manually painted through a dedicated user interface [Ion et al. 2016] or automatically computed through, e.g. topology optimization [Zhu et al. 2017].

In this paper we focus on defining a novel class of parametric metamaterials. A key challenge – besides triggering the desired elastic behavior – is to enforce constraints due to the additive manufacturing processes. Most existing works on the topic have focused on processes such as selective laser sintering (SLS) and stereolithography (SLA). These technologies have different constraints than the widely available and inexpensive fused filament fabrication (FFF) 3D printers. In particular, most recent works produce geometries exhibiting truss structures (detailed in Section 2). After slicing, these structures produce many isolated ellipses in each slice. These are extremely challenging to print on FFF systems, where continuity

# Stitch Meshing

KUI WU, University of Utah

XIFENG GAO, New York University

ZACHARY FERGUSON, New York University

DANIELE PANOZZO, New York University

CEM YUKSEL, University of Utah



**Fig. 1.** Example yarn-level models generated from input 3D surfaces using our fully automatic pipeline.

We introduce the first fully automatic pipeline to convert arbitrary 3D shapes into knit models. Our pipeline is based on a global parametrization remeshing pipeline to produce an isotropic quad-dominant mesh aligned with a 2-RoS field. The knitting directions over the surface are determined using a set of custom topological operations and a two-step global optimization that minimizes the number of irregularities. The resulting mesh is converted into a valid stitch mesh that represents the knit model. The yarn curves are generated from the stitch mesh and the final yarn geometry is computed using a yarn-level relaxation process. Thus, we produce topologically valid models that can be used with a yarn-level simulation. We validate our algorithm by automatically generating knit models from complex 3D shapes and processing over a hundred models with various shapes without any user input or parameter tuning. We also demonstrate applications of our approach for custom knit model generation using fabrication via 3D printing.

**CCS Concepts:** • **Computing methodologies** → **Mesh geometry models**;

Authors' addresses: Kui Wu, University of Utah, [kwu@cs.utah.edu](mailto:kwu@cs.utah.edu); Xifeng Gao, New York University, [gxf.xisha@gmail.com](mailto:gxf.xisha@gmail.com); Zachary Ferguson, New York University, [zfergus@nyu.edu](mailto:zfergus@nyu.edu); Daniele Panozzo, New York University, [panozzo@nyu.edu](mailto:panozzo@nyu.edu); Cem Yuksel, University of Utah, [cem@cemyuksel.com](mailto:cem@cemyuksel.com).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART130 \$15.00

<https://doi.org/10.1145/3197517.3201360>

Additional Key Words and Phrases: Yarn-level cloth, yarn-level modeling

## ACM Reference Format:

Kui Wu, Xifeng Gao, Zachary Ferguson, Daniele Panozzo, and Cem Yuksel. 2018. Stitch Meshing. *ACM Trans. Graph.* 37, 4, Article 130 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201360>

## 1 INTRODUCTION

Knitted garments are common in our daily lives, going from socks and T-shirts to winter clothing and accessories, and are thus ubiquitous in movies and games. There are two good reasons for favoring knitting over its alternatives: knitted fabrics easily stretch and the shaping techniques used in knitting allow producing complex 3D surfaces without any seams.

However, designing knitting pattern for a given 3D surface is still an open problem. Knitting patterns are currently designed using a high level of expertise and numerous iterations of trial and error to figure out how one could knit a particular 3D shape. That is why most knitting patterns used today are merely derivations of a limited number of well-known and well-understood shapes. In computer graphics, stitch meshes [Yuksel et al. 2012] provide a powerful interface for modeling knit garments. However, they still require users to manually design the topology of the given (typically low-resolution) input mesh. This requires the user to know exactly how to knit the desired shape and prepare an input mesh accordingly. Therefore, it is extremely difficult and time consuming to design knitted models for complex and uncommon shapes, like the ones

# Interactive Exploration of Design Trade-Offs

ADRIANA SCHULZ and HARRISON WANG, Massachusetts Institute of Technology, USA

EITAN GRINSUN, Columbia University, USA

JUSTIN SOLOMON and WOJCIECH MATUSIK, Massachusetts Institute of Technology, USA

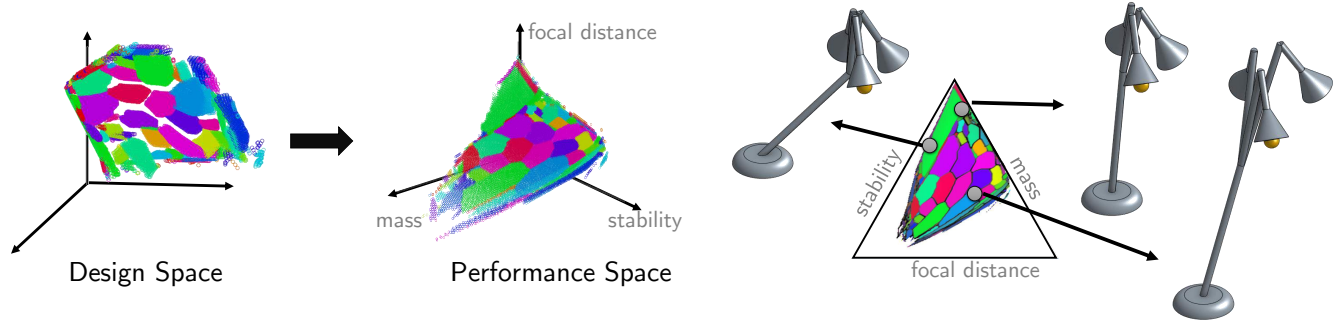


Fig. 1. Our method allows users to optimize designs based on a set of performance metrics. Given a design space and a set of performance evaluation functions, our method automatically extracts the *Pareto set*—those design points with optimal trade-offs. We represent Pareto points in design and performance space with a set of corresponding manifolds (left). The Pareto-optimal solutions are then embedded to allow interactive exploration of performance trade-offs (right). The mapping from manifolds in performance space back to design space allows designers to explore performance trade-offs interactively while visualizing the corresponding geometry and gaining an understanding of a model’s underlying properties.

Typical design for manufacturing applications requires simultaneous optimization of conflicting performance objectives: Design variations that improve one performance metric may decrease another performance metric. In these scenarios, there is no unique optimal design but rather a set of designs that are optimal for different trade-offs (called Pareto-optimal). In this work, we propose a novel approach to discover the Pareto front, allowing designers to navigate the landscape of compromises efficiently. Our approach is based on a first-order approximation of the Pareto front, which allows entire neighborhoods rather than individual points on the Pareto front to be captured. In addition to allowing for efficient discovery of the Pareto front and the corresponding mapping to the design space, this approach allows us to represent the entire trade-off manifold as a small collection of patches that comprise a high-quality and piecewise-smooth approximation. We illustrate how this technique can be used for navigating performance trade-offs in computer-aided design (CAD) models.

CCS Concepts: • **Computing methodologies** → **Shape modeling**; *Graphics systems and interfaces*;

Additional Key Words and Phrases: Pareto optimality, design for manufacturing, shape modeling

Authors’ addresses: Adriana Schulz, [aschulz@csail.mit.edu](mailto:aschulz@csail.mit.edu); Harrison Wang, [hwang123@mit.edu](mailto:hwang123@mit.edu), Massachusetts Institute of Technology, Cambridge, MA, 02138, USA; Eitan Grinspun, [eitan@cs.columbia.edu](mailto:eitan@cs.columbia.edu), Columbia University, New York, NY, 10027, USA; Justin Solomon, [jsolomon@mit.edu](mailto:jsolomon@mit.edu); Wojciech Matusik, [wojciech@csail.mit.edu](mailto:wojciech@csail.mit.edu), Massachusetts Institute of Technology, Cambridge, MA, 02139, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART131 \$15.00

<https://doi.org/10.1145/3197517.3201385>

## ACM Reference Format:

Adriana Schulz, Harrison Wang, Eitan Grinspun, Justin Solomon, and Wojciech Matusik. 2018. Interactive Exploration of Design Trade-Offs. *ACM Trans. Graph.* 37, 4, Article 131 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201385>

## 1 INTRODUCTION

In typical manufacturing applications, design goals are measured by how the finished products *perform* in the physical world. Design is driven by objectives such as weight, stability, durability, compliance, and other functionality-related metrics. Directly specifying designs that optimize performance objectives is challenging since it requires predicting physical behavior. For this reason, performance-driven computational methods are critical and have been widely explored in recent research on manufacturing-oriented design. These approaches guide users in efficiently exploring design spaces while optimizing for a given performance objective, which can be measured using physical simulations.

A fundamental limitation of typical design optimization techniques is that they require a single objective function for evaluating performance. In most applications, however, multiple criteria are used to evaluate the quality of a design. Structures must be stable *and* lightweight. Vehicles must be aerodynamic, durable, *and* inexpensive to produce. In most cases, the performance objectives are not only multiple but also *conflicting*: improving a design on one axis often decreases its quality on another axis. In reality, designers and engineers navigate a complex landscape of compromises, generating objects that perhaps do not optimize any single quality measure but rather are considered optimal under a given performance trade-off.

Since it is impossible to optimize more than one criterion at a time, standard optimization approaches require expressing a set of performance criteria in a combined objective function that balances



# Autocomplete 3D Sculpting

MENGQI PENG, University of Hong Kong

JUN XING, University of Hong Kong and USC Institute for Creative Technologies

LI-YI WEI, University of Hong Kong and Adobe Research

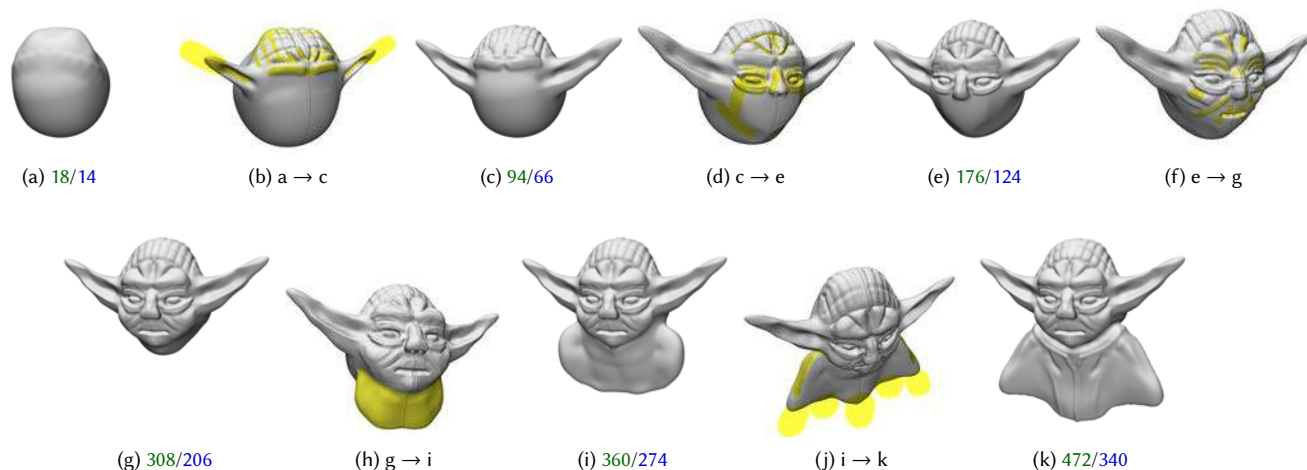


Fig. 1. Our *autocomplete 3D sculpting system*. With an authoring interface similar to common sculpting tools, our system detects potential repetitions and provides suggestions for which the users can accept, partially accept, or ignore. Shown here is an editing sequence with the number of cumulative **manual** and **autocomplete** strokes indicated below each stage, and in-between steps with some hints visualized in yellow colors, including overlapped/non-overlapped strokes for both surface and freeform types. Other functionalities include workflow clone and automatic camera view control. Please refer to the accompanying videos for recorded actions, including all edits, hints, and camera views.

Digital sculpting is a popular means to create 3D models but remains a challenging task. We propose a 3D sculpting system that assists users, especially novices, in freely creating models with reduced input labor and enhanced output quality. With an interactive sculpting interface, our system silently records and analyzes users' workflows including brush strokes and camera movements, and predicts what they might do in the future. Users can accept, partially accept, or ignore the suggestions and thus retain full control and individual style. They can also explicitly select and clone past workflows over output model regions. Our key idea is to consider *how* a model is authored via dynamic workflows in addition to *what* is shaped in static geometry. This allows our method for more accurate analysis of user intentions and more general synthesis of shape structures than prior workflow or geometry methods, such as large overlapping deformations. We evaluate our method via user feedbacks and authored models.

Authors' addresses: Mengqi Peng, University of Hong Kong; Jun Xing, University of Hong Kong, USC Institute for Creative Technologies; Li-Yi Wei, University of Hong Kong, Adobe Research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART132 \$15.00 <https://doi.org/10.1145/3197517.3201297>

CCS Concepts: • **Human-centered computing** → **User interface design**; • **Computing methodologies** → **Shape modeling**;

Additional Key Words and Phrases: workflow, autocomplete, clone, beautification, sculpting, modeling, user interface

## ACM Reference Format:

Mengqi Peng, Jun Xing, and Li-Yi Wei. 2018. Autocomplete 3D Sculpting. *ACM Trans. Graph.* 37, 4, Article 132 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201297>

## 1 INTRODUCTION

Digital sculpting is a popular means to create 3D models with diverse styles and organic shapes. The authoring practices often involves repetitive applications of large, overlapping deformations interleaved with small, detailed strokes. Such cumulative process can demand significant expertise and efforts, and be particularly daunting for novice users.

Significant research has been devoted to analyzing and synthesizing 3D geometry based on shapes [Chaudhuri and Koltun 2010; Funkhouser et al. 2004; Mitra et al. 2013; Takayama et al. 2011] or procedures [Emilien et al. 2015; Nishida et al. 2016]. These methods mainly focus on *what* the model is shaped instead of *how* it is authored, and thus cannot readily apply to iterative, dynamic sculpting of general shapes. The processes of how models are authored by users, termed workflows, contain rich information that

# FoldSketch: Enriching Garments with Physically Reproducible Folds

MINCHEN LI, The University of British Columbia  
ALLA SHEFFER, The University of British Columbia  
EITAN GRINSPUN, Columbia University  
NICHOLAS VINING, The University of British Columbia

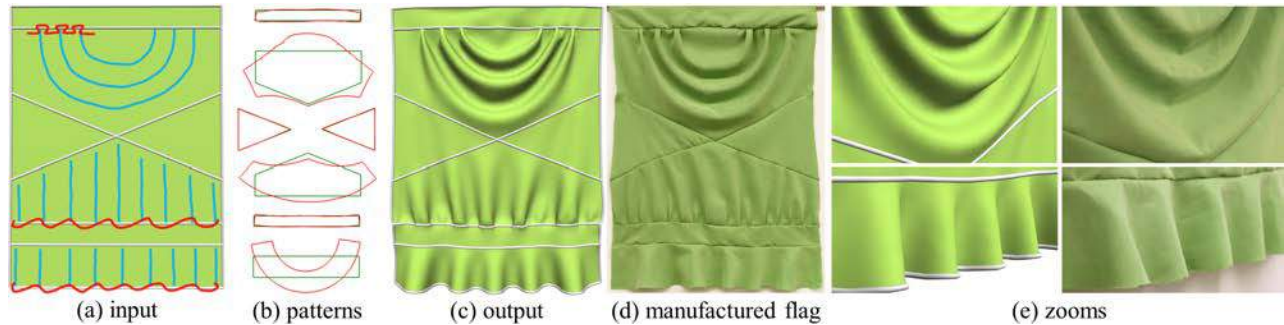


Fig. 1. Complex physically simulated fabric folds generated using FoldSketch: (left to right) Plain input flag with user sketched schematic folds, original (green) and modified final (red) patterns; final post-simulation flag augmented with user-expected folds; real-life replica manufactured using the produced patterns; zooming in highlights the complex and evolving output fold profile shapes.

While folds and pleats add interest to garments and cloth objects, incorporating them into an existing design manually or using existing software requires expertise and time. We present *FoldSketch*, a new system that supports simple and intuitive fold and pleat design. FoldSketch users specify the fold or pleat configuration they seek using a simple schematic sketching interface; the system then algorithmically generates both the fold-enhanced 3D garment geometry that conforms to user specifications, and the corresponding 2D patterns that reproduce this geometry within a simulation engine. While previous work aspired to compute the desired patterns for a given *target* 3D garment geometry, our main algorithmic challenge is that we do not have target geometry to start with. Real-life garment folds have complex profile shapes, and their exact geometry and location on a garment are intricately linked to a range of physical factors such as fabric properties and the garment's interaction with the wearer's body; it is therefore virtually impossible to predict the 3D shape of a fold-enhanced garment using purely geometric means. At the same time, using physical simulation to model folds requires appropriate 2D patterns and initial drape, neither of which can be easily provided by the user. We obtain both the 3D fold-enhanced garment and its corresponding patterns and initial drape via an alternating 2D-3D algorithm. We first expand the input patterns by allocating excess material for the expected fold formation; we then use these patterns to produce an estimated fold-enhanced drape geometry that balances designer expectations against physical reproducibility. We use the patterns and the estimated drape as input to a simulation generating an initial reproducible output. We improve

the output's alignment with designer expectations by progressively refining the patterns and the estimated drape, converging to a final fully physically reproducible fold-enhanced garment. Our experiments confirm that FoldSketch reliably converges to a desired garment geometry and corresponding patterns and drape, and works well with different physical simulators. We demonstrate the versatility of our approach by showcasing a collection of garments augmented with diverse fold and pleat layouts specified via the FoldSketch interface, and further validate our approach via comparisons to alternative solutions and feedback from potential users.

CCS Concepts: • **Computing methodologies** → **Mesh geometry models**; *Physical simulation*;

Additional Key Words and Phrases: garment modeling, sketch-based modeling, folds

## ACM Reference Format:

Minchen Li, Alla Sheffer, Eitan Grinspun, and Nicholas Vining. 2018. FoldSketch: Enriching Garments with Physically Reproducible Folds. *ACM Trans. Graph.* 37, 4, Article 133 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201310>

## 1 INTRODUCTION

Fashion designers frequently use strategically placed folds and pleats to add interest to garments and other cloth objects to increase their visual appeal. Pleats and folds are typically formed by gathering fabric along a seamline and stitching the gathered fabric to hold it in place. In both traditional garment design software and manual workflows, these features are incorporated into an existing design by first skillfully modifying the underlying 2D patterns, and then *draping* the garment atop a mannequin or a dress form by meticulously positioning it to achieve the desired fold look. This workflow requires expertise and time, since it often takes multiple trial and error iterations to successfully incorporate the envisioned folds or pleats into an existing design [Arnold 1985; Kiisel 2013]. We propose *FoldSketch*, a new algorithmic framework for fold and pleat

Authors' addresses: Minchen Li, The University of British Columbia, minchernl@gmail.com; Alla Sheffer, The University of British Columbia, sheffa@cs.ubc.ca; Eitan Grinspun, Columbia University, eitan@cs.columbia.edu; Nicholas Vining, The University of British Columbia, nvining@cs.ubc.ca.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
0730-0301/2018/8-ART133 \$15.00  
<https://doi.org/10.1145/3197517.3201310>

# Support-Free Volume Printing by Multi-Axis Motion

CHENGKAI DAI, Delft University of Technology, The Netherlands

CHARLIE C. L. WANG\*, Delft University of Technology, The Netherlands

CHENMING WU, Tsinghua University, China

SYLVAIN LEFEBVRE, INRIA, France

GUOXIN FANG, Delft University of Technology, The Netherlands

YONG-JIN LIU, Tsinghua University, China



Fig. 1. Our method enables support-free 3D printing of solid models. By exploiting all 6 degrees of freedom (translations, rotations) and depositing material along *curved* layers, we make support structures unnecessary in most cases. This increases further the flexibility offered by 3D printing, such as freeing designers from support constraints on complex parts.

This paper presents a new method to fabricate 3D models on a robotic printing system equipped with multi-axis motion. Materials are accumulated inside the volume along curved tool-paths so that the need of supporting structures can be tremendously reduced – if not completely abandoned – on all models. Our strategy to tackle the challenge of tool-path planning for multi-axis 3D printing is to perform two successive decompositions, first volume-to-surfaces and then surfaces-to-curves. The volume-to-surfaces decomposition is achieved by optimizing a scalar field within the volume that represents the fabrication sequence. The field is constrained such that its iso-values represent curved layers that are supported from below, and present a convex surface affording for collision-free navigation of the printer head. After extracting all curved layers, the surfaces-to-curves decomposition

covers them with tool-paths while taking into account constraints from the robotic printing system. Our method successfully generates tool-paths for 3D printing models with large overhangs and high-genus topology. We fabricated several challenging cases on our robotic platform to verify and demonstrate its capabilities.

CCS Concepts: • **Computing methodologies** → **Shape modeling**;

Additional Key Words and Phrases: 3D printing, multi-axis motion, support-free structures, tool-path generation

## ACM Reference Format:

Chengkai Dai, Charlie C. L. Wang, Chenming Wu, Sylvain Lefebvre, Guoxin Fang, and Yong-jin Liu. 2018. Support-Free Volume Printing by Multi-Axis Motion. *ACM Trans. Graph.* 37, 4, Article 1 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201342>

## 1 INTRODUCTION

The prompt development of *additive manufacturing* (AM) techniques has motivated significant research effort in the area of computer graphics, computer-aided design, biomedical engineering and robotics (e.g., [Liu et al. 2014; Shamir et al. 2016]). Although it is called 3D printing, the fabrication in most commercial systems is still taken in a 2.5D manner – materials are accumulated layer upon layer in planes along a fixed printing direction. This significantly reduces the complexity and development cost of both hardware and software. However, this introduces additional manufacturability constraints,

\*Corresponding author: c.c.wang@tudelft.nl (Charlie C. L. Wang)

Authors' addresses: Chengkai Dai, Delft University of Technology, Delft, The Netherlands; Charlie C. L. Wang, Delft University of Technology, Delft, The Netherlands; Chenming Wu, Tsinghua University, Beijing, China; Sylvain Lefebvre, INRIA, France; Guoxin Fang, Delft University of Technology, Delft, The Netherlands; Yong-jin Liu, Tsinghua University, Beijing, China.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART1 \$15.00

<https://doi.org/10.1145/3197517.3201342>



# CoreCavity: Interactive Shell Decomposition for Fabrication with Two-Piece Rigid Molds

KAZUTAKA NAKASHIMA, The University of Tokyo

THOMAS AUZINGER, IST Austria

EMMANUEL IARUSSI, CONICET, IST Austria

RAN ZHANG, IST Austria

TAKEO IGARASHI, The University of Tokyo

BERND BICKEL, IST Austria

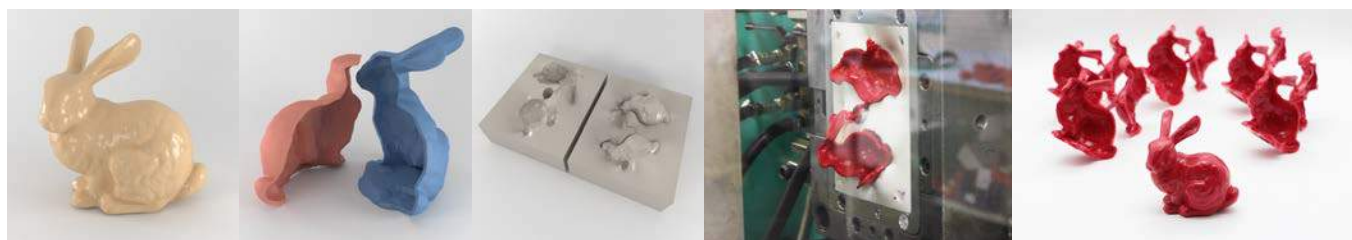


Fig. 1. From left to right: We propose a method to decompose an input shape (surface mesh) into molding-ready shell pieces and to generate the corresponding 3D printable molds. Using, for example, an injection molding machine, we can fabricate many physical replica at low cost and in a short amount of time.

Molding is a popular mass production method, in which the initial expenses for the mold are offset by the low per-unit production cost. However, the physical fabrication constraints of the molding technique commonly restrict the shape of moldable objects. For a complex shape, a decomposition of the object into moldable parts is a common strategy to address these constraints, with plastic model kits being a popular and illustrative example. However, conducting such a decomposition requires considerable expertise, and it depends on the technical aspects of the fabrication technique, as well as aesthetic considerations. We present an interactive technique to create such decompositions for two-piece molding, in which each part of the object is cast between two rigid mold pieces. Given the surface description of an object, we decompose its thin-shell equivalent into moldable parts by first performing a coarse decomposition and then utilizing an active contour model for the boundaries between individual parts. Formulated as an optimization problem, the movement of the contours is guided by an energy reflecting fabrication constraints to ensure the *moldability* of each part. Simultaneously, the user is provided with editing capabilities to enforce aesthetic guidelines. Our interactive interface provides control of the contour positions by allowing, for example, the alignment of part boundaries with object features. Our technique enables a novel workflow, as it empowers novice users to explore the design space, and it generates fabrication-ready two-piece molds that

Authors' addresses: Kazutaka Nakashima, The University of Tokyo, taka@ui.is.s.u-tokyo.ac.jp; Thomas Auzinger, IST Austria, thomas.auzinger@ist.ac.at; Emmanuel Iarussi, CONICET, IST Austria, eiarussi@conicet.gov.ar; Ran Zhang, IST Austria, rzhang@ist.ac.at; Takeo Igarashi, The University of Tokyo, takeo@acm.org; Bernd Bickel, IST Austria, bernd.bickel@ist.ac.at.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART135 \$15.00

<https://doi.org/10.1145/3197517.3201341>

can be used either for casting or industrial injection molding of free-form objects.

CCS Concepts: • **Applied computing** → **Computer-aided manufacturing**; • **Computing methodologies** → *Shape analysis*; • **Mathematics of computing** → *Mesh generation*; • **Human-centered computing** → *Graphical user interfaces*;

Additional Key Words and Phrases: molding, fabrication, height field, decomposition

## ACM Reference Format:

Kazutaka Nakashima, Thomas Auzinger, Emmanuel Iarussi, Ran Zhang, Takeo Igarashi, and Bernd Bickel. 2018. CoreCavity: Interactive Shell Decomposition for Fabrication with Two-Piece Rigid Molds. *ACM Trans. Graph.* 37, 4, Article 135 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201341>

## 1 INTRODUCTION

In recent years, the rise of 3D printing technology has democratized the fabrication of complex shapes. Nevertheless, classical manufacturing approaches, such as molding, are often the preferred choice because of two main reasons: they scale extremely well with the amount of required copies, and they draw from a wide range of available base materials. Unfortunately, in contrast to 3D printing, which seemingly requires only a few clicks to replicate an object, obtaining a mold is much more challenging.

In this paper, we focus on designing re-usable rigid molds for *two-piece molding*, the simplest and most cost-efficient variant of molding. In two-piece molding, as illustrated in Fig. 2, two rigid mold pieces form a cavity filled with a liquid or pliable material, i.e., resin, plastic, metal, or ceramics, which then solidifies. Subsequently, the mold pieces are separated, and the cast object is released by a linear translation along a so-called *parting direction* (PD). To make the

# Metamolds: Computational Design of Silicone Molds

THOMAS ALDERIGHI, University of Pisa and ISTI - CNR  
LUIGI MALOMO, ISTI - CNR  
DANIELA GIORGI, ISTI - CNR  
NICO PIETRONI, University of Technology Sidney and ISTI - CNR  
BERND BICKEL, IST Austria  
PAOLO CIGNONI, ISTI - CNR



Fig. 1. Our approach automatically generates the shape of silicone molds for casting an input 3D mesh into moldable pieces (left), we design and 3D print a set of metamolds (middle, in red). The actual mold pieces (right, in green and whitish) are produced by pouring liquid silicone onto the metamolds, then assembled and used to cast copies of the input model.

We propose a new method for fabricating digital objects through reusable silicone molds. Molds are generated by casting liquid silicone into custom 3D printed containers called *metamolds*. Metamolds automatically define the cuts that are needed to extract the cast object from the silicone mold. The shape of metamolds is designed through a novel segmentation technique, which takes into account both geometric and topological constraints involved in the process of mold casting. Our technique is simple, does not require changing the shape or topology of the input objects, and only requires off-the-shelf materials and technologies. We successfully tested our method on a set of challenging examples with complex shapes and rich geometric detail.

CCS Concepts: • **Computing methodologies** → *Shape modeling*;

Additional Key Words and Phrases: fabrication, molding, casting

## ACM Reference Format:

Thomas Alderighi, Luigi Malomo, Daniela Giorgi, Nico Pietroni, Bernd Bickel, and Paolo Cignoni. 2018. Metamolds: Computational Design of Silicone Molds. *ACM Trans. Graph.* 37, 4, Article 136 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201381>

Authors' addresses: Thomas Alderighi, University of Pisa, ISTI - CNR, [thomas.alderighi@isti.cnr.it](mailto:thomas.alderighi@isti.cnr.it); Luigi Malomo, ISTI - CNR; Daniela Giorgi, ISTI - CNR; Nico Pietroni, University of Technology Sidney, ISTI - CNR; Bernd Bickel, IST Austria; Paolo Cignoni, ISTI - CNR.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART136 \$15.00

<https://doi.org/10.1145/3197517.3201381>

## 1 INTRODUCTION

While 3D printing technologies are becoming faster and more precise, classical manufacturing techniques remain the first choice for most industrial application scenarios. Industrial production is still largely dominated by casting techniques: casting scales well with the number of copies, supports a wide spectrum of materials, and ensures high geometric accuracy.

A popular casting technique for high-quality reproduction of art objects is silicone mold casting [Bruckner et al. 2010]. For simple scenarios, a physical prototype is submerged in liquid silicone; the cured silicone forms a mold around the object; then, the prototype is extracted by manually cutting and opening the silicone mold. Multiple copies can be cast by filling the silicone mold with a liquid casting material such as resin. Silicone molding has two main practical advantages over traditional rigid casting: the replicas can be safely extracted by deforming the flexible mold without damaging it, and overhanging geometric details do not constitute a severe limitation.

While conceptually simple, silicone mold casting may become extremely challenging when applied to non-trivial shapes, and often requires the intervention of skilled professionals. For example, objects with handles usually need a set of carefully placed extra cuts to make the extraction physically possible. Moreover, venting pipes have to be attached to the prototype object before submersion in liquid silicone, to let the air flow out and avoid artifacts in the replicas due to trapped air bubbles.

Recently, some methods tried to reinterpret the manufacturing process of mold casting using a computational approach [Babaei et al. 2017; Herholz et al. 2015; Malomo et al. 2016]. However, none

# DSCarver: Decompose-and-Spiral-Carve for Subtractive Manufacturing

HAISEN ZHAO, Shandong University

HAO ZHANG, Simon Fraser University

SHIQING XIN, Shandong University

YUANMIN DENG, Shandong University

CHANGHE TU, Shandong University

WENPING WANG, Hong Kong University

DANIEL COHEN-OR, Tel Aviv University

BAOQUAN CHEN, Shandong University and Peking University

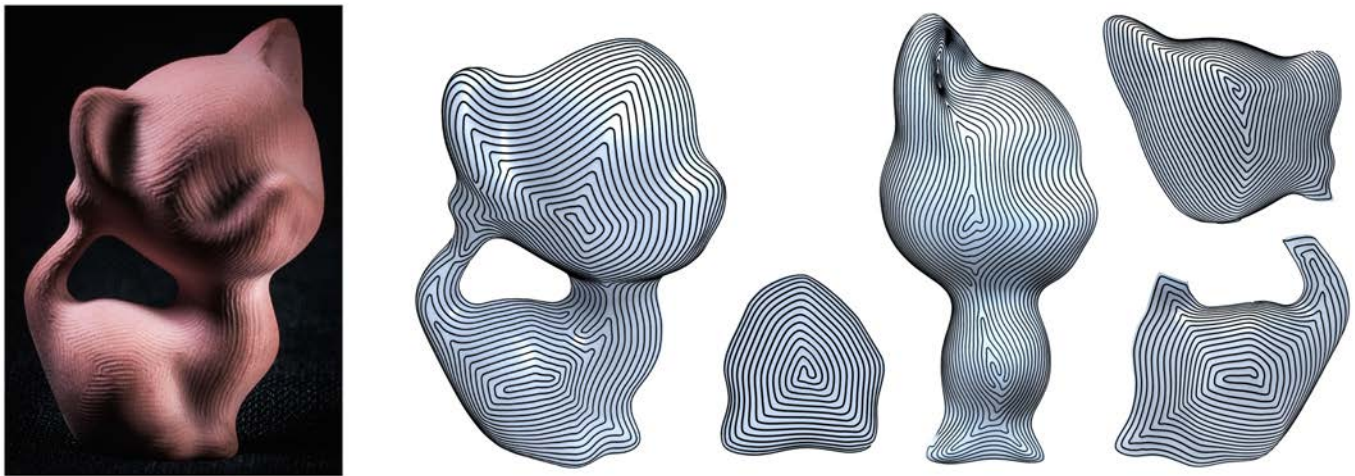


Fig. 1. A freeform 3D object, the KITTEN, automatically decomposed into five machinable patches, each carved using a CNC machine following a continuous iso-scallop Fermat spiral path. The fully machined object is shown on the left. We deliberately lowered the path resolution to ease path visualization.

We present an automatic algorithm for *subtractive* manufacturing of freeform 3D objects using high-speed machining (HSM) via CNC. A CNC machine operates a cylindrical cutter to carve off material from a 3D shape stock, following a tool path, to “expose” the target object. Our method decomposes the input object’s surface into a small number of patches each of which is fully accessible and machinable by the CNC machine, in *continuous* fashion, under a fixed cutter-object setup configuration. This is achieved by covering the input surface with a minimum number of *accessible regions* and then extracting a set of machinable patches from each accessible region. For

Authors’ addresses: Haisen Zhao, Shandong University; Hao Zhang, Simon Fraser University; Shiqing Xin, Shandong University; Yuanmin Deng, Shandong University; Changhe Tu, Shandong University; Wenping Wang, Hong Kong University; Daniel Cohen-Or, Tel Aviv University; Baoquan Chen, Shandong University and Peking University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART137 \$15.00

<https://doi.org/10.1145/3197517.3201338>

each patch obtained, we compute a continuous, space-filling, and *iso-scallop* tool path which conforms to the patch boundary, enabling efficient carving with high-quality surface finishing. The tool path is generated in the form of *connected Fermat spirals*, which have been generalized from a 2D fill pattern for layered manufacturing to work for curved surfaces. Furthermore, we develop a novel method to control the spacing of Fermat spirals based on directional surface curvature and adapt the heat method to obtain iso-scallop carving. We demonstrate automatic generation of accessible and machinable surface decompositions and iso-scallop Fermat spiral carving paths for freeform 3D objects. Comparisons are made to tool paths generated by commercial software in terms of real machining time and surface quality.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; *Shape analysis*;

Additional Key Words and Phrases: Subtractive manufacturing, tool path planning, CNC, surface decomposition, set cover, Fermat spirals

## ACM Reference Format:

Haisen Zhao, Hao Zhang, Shiqing Xin, Yuanmin Deng, Changhe Tu, Wenping Wang, Daniel Cohen-Or, and Baoquan Chen. 2018. DSCarver: Decompose-and-Spiral-Carve for Subtractive Manufacturing. *ACM Trans. Graph.* 37, 4, Article 137 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201338>



# Scanner: Efficient Video Analysis at Scale

ALEX POMS, Carnegie Mellon University, USA

WILL CRICHTON, PAT HANRAHAN, and KAYVON FATAHALIAN, Stanford University, USA

A growing number of visual computing applications depend on the analysis of large video collections. The challenge is that scaling applications to operate on these datasets requires efficient systems for pixel data access and parallel processing across large numbers of machines. Few programmers have the capability to operate efficiently at these scales, limiting the field's ability to explore new applications that leverage big video data. In response, we have created Scanner, a system for productive and efficient video analysis at scale. Scanner organizes video collections as tables in a data store optimized for sampling frames from compressed video, and executes pixel processing computations, expressed as dataflow graphs, on these frames. Scanner schedules video analysis applications expressed using these abstractions onto heterogeneous throughput computing hardware, such as multi-core CPUs, GPUs, and media processing ASICs, for high-throughput pixel processing. We demonstrate the productivity of Scanner by authoring a variety of video processing applications including the synthesis of stereo VR video streams from multi-camera rigs, markerless 3D human pose reconstruction from video, and data-mining big video datasets such as hundreds of feature-length films or over 70,000 hours of TV news. These applications achieve near-expert performance on a single machine and scale efficiently to hundreds of machines, enabling formerly long-running big video data analysis tasks to be carried out in minutes to hours.

CCS Concepts: • **Computing methodologies** → *Graphics systems and interfaces*; Image processing;

Additional Key Words and Phrases: large-scale video processing

## ACM Reference Format:

Alex Poms, Will Crichton, Pat Hanrahan, and Kayvon Fatahalian. 2018. Scanner: Efficient Video Analysis at Scale. *ACM Trans. Graph.* 37, 4, Article 138 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201394>

## 1 INTRODUCTION

The world is increasingly instrumented with sources of video: cameras are commonplace on people (smartphone cameras, GoPros), on vehicles (automotive cameras, drone videography), and in urban environments (traffic cameras, security cameras). Extracting value from these high-resolution video streams is a key research and commercial challenge, and a growing number of applications in fields like computer graphics, vision, robotics and basic science are based on analyzing large amounts of video.

The challenge is that scaling video analysis tasks to large video collections (thousands of hours of cable TV or YouTube clips, the output of a modern VR video capture rig) requires optimized systems

for managing pixel data as well as efficient, parallel processing on accelerated computing hardware (clusters of multi-core CPUs, GPUs, and ASICs). Unfortunately, very few programmers have the skill set to implement efficient software for processing large video datasets, inhibiting the field's ability to explore new applications that leverage this data. Inspired by the impact of data analytics frameworks such as MapReduce [Dean and Ghemawat 2004] and Spark [Zaharia et al. 2010], which facilitate rapid development of scalable big-data analytics applications, we have created *Scanner*, a system for productive and efficient *big video data* analysis.

Scanner provides integrated system support for two performance-critical aspects of video analysis: storing and accessing pixel data from large video collections, and executing expensive pixel-level operations in parallel on large numbers of video frames. Scanner addresses the first need by organizing video collections and derived raster data (depth maps, activation maps, flow fields, etc.) as tables in a data store whose implementation is optimized for compressed video. It addresses the second need by organizing pixel-analysis tasks as dataflow graphs that operate on sequences of frames sampled from tables. Scanner graphs support features useful for video processing, such as sparse sampling of video frames, access to temporal windows of frames, and state propagation across computations on successive frames. Scanner schedules these computations efficiently onto heterogeneous computing hardware such as multi-core CPUs, GPUs, and media processing ASICs.

We demonstrate that applications using Scanner for expensive, pixel-level video processing operations achieve near-expert performance when deployed on workstations with high-core count CPUs and multiple GPUs. The same applications also scale efficiently to hundreds of machines without source-level change. We report on experiences using Scanner to implement several large-scale video analysis applications including VR video processing, 3D human pose reconstruction from multi-viewpoint video, and data mining large video datasets of TV news. In these cases, Scanner enabled video analysis tasks that previously required days of processing (when implemented by researchers and data scientists using ad hoc solutions) to be carried out efficiently in hours to minutes.

## 2 CHALLENGES OF VIDEO ANALYSIS

Executing pixel-analysis pipelines (e.g., feature extraction, face/object detection, image similarity and alignment) on large image collections is the performance-critical component of many big visual data applications such as data-driven image manipulation and enhancement [Hays and Efros 2007; Kemelmacher-Shlizerman 2016], novel techniques for organizing and browsing photo collections [Sivic et al. 2008; Snavely et al. 2006], and exploratory data mining of the visual world [Chen et al. 2013; Doersch et al. 2012; Ginosar et al. 2017; Matzen et al. 2017; Zhu et al. 2014]. While these early applications analyzed collections of images, a growing class of applications now seek to manipulate large video datasets. To better understand

---

Authors' addresses: Alex Poms, Carnegie Mellon University, USA; Will Crichton; Pat Hanrahan; Kayvon Fatahalian, Stanford University, USA.

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART138 \$15.00

<https://doi.org/10.1145/3197517.3201394>

# Differentiable Programming for Image Processing and Deep Learning in Halide

TZU-MAO LI, MIT CSAIL

MICHAËL GHARBI, MIT CSAIL

ANDREW ADAMS, Facebook AI Research

FRÉDO DURAND, MIT CSAIL

JONATHAN RAGAN-KELLEY, UC Berkeley & Google

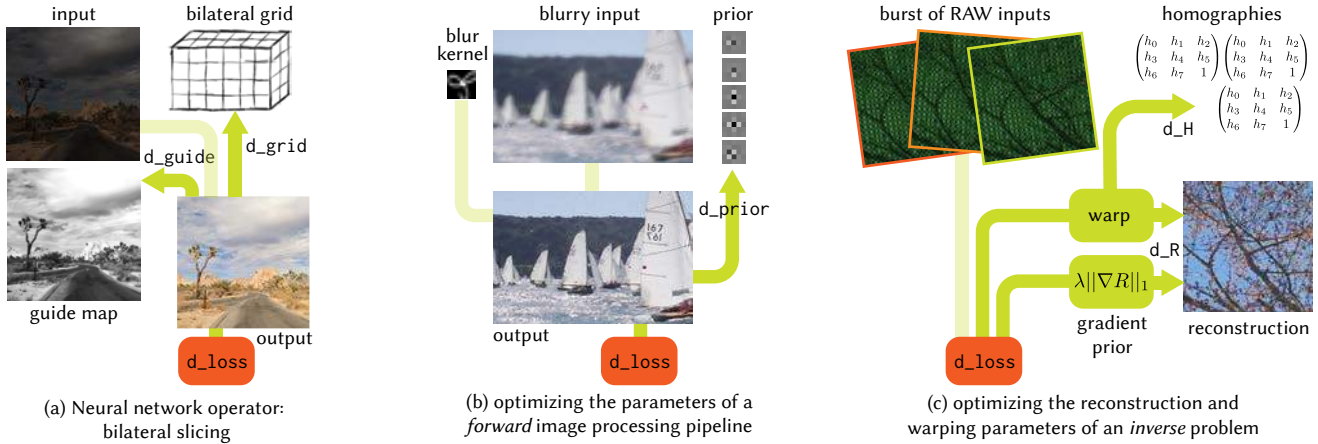


Fig. 1. Our system automatically derives and optimizes gradient code for general image processing pipelines, and yields state-of-the-art performance on both CPUs and GPUs. This enables a variety of imaging applications, from training novel neural network layers (a), to optimizing the parameters of traditional image processing pipelines (b), to solving inverse reconstruction problems (c). To support these applications, we extend the Halide language with features to automatically and efficiently compute gradients of arbitrary programs. We also introduce a new automatic performance optimization that can handle the specific computation patterns of gradient computation. Using our system, a user can easily write high-level image processing algorithms, and then automatically derive high-performance gradient code for CPUs, GPUs, and other architectures. Images from left to right are from MIT5k dataset [Bychkovsky et al. 2011], ImageNet [Deng et al. 2009], and deep demosaicking dataset [Gharbi et al. 2016], respectively.

Gradient-based optimization has enabled dramatic advances in computational imaging through techniques like deep learning and nonlinear optimization. These methods require gradients not just of simple mathematical functions, but of general *programs* which encode complex transformations of images and graphical data. Unfortunately, practitioners have traditionally been limited to either hand-deriving gradients of complex computations, or composing programs from a limited set of coarse-grained operators in deep learning frameworks. At the same time, writing programs with the level of performance needed for imaging and deep learning is prohibitively difficult for most programmers.

We extend the image processing language Halide with general reverse-mode automatic differentiation (AD), and the ability to automatically optimize the implementation of gradient computations. This enables automatic

computation of the gradients of arbitrary Halide programs, at high performance, with little programmer effort. A key challenge is to structure the gradient code to retain parallelism. We define a simple algorithm to automatically schedule these pipelines, and show how Halide’s existing scheduling primitives can express and extend the key AD optimization of “checkpointing.”

Using this new tool, we show how to easily define new neural network layers which automatically compile to high-performance GPU implementations, and how to solve nonlinear inverse problems from computational imaging. Finally, we show how differentiable programming enables dramatically improving the quality of even traditional, feed-forward image processing algorithms, blurring the distinction between classical and deep methods.

CCS Concepts: • **Computing methodologies** → **Graphics systems and interfaces**; *Machine learning*; *Image processing*;

Additional Key Words and Phrases: image processing, deep learning, automatic differentiation

## ACM Reference Format:

Tzu-Mao Li, Michaël Gharbi, Andrew Adams, Frédo Durand, and Jonathan Ragan-Kelley. 2018. Differentiable Programming for Image Processing and Deep Learning in Halide. *ACM Trans. Graph.* 37, 4, Article 139 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201383>

Authors’ addresses: Tzu-Mao Li, MIT CSAIL, tzumao@mit.edu; Michaël Gharbi, MIT CSAIL, gharbi@mit.edu; Andrew Adams, Facebook AI Research, andrew.b.adams@gmail.com; Frédo Durand, MIT CSAIL, fred@mit.edu; Jonathan Ragan-Kelley, UC Berkeley & Google, jrk@berkeley.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART139

<https://doi.org/10.1145/3197517.3201383>

# A High-Performance Software Graphics Pipeline Architecture for the GPU

MICHAEL KENZEL, BERNHARD KERBL, and DIETER SCHMALSTIEG, Graz University of Technology, Austria  
MARKUS STEINBERGER, Graz University of Technology, Austria and Max Planck Institute for Informatics, Germany

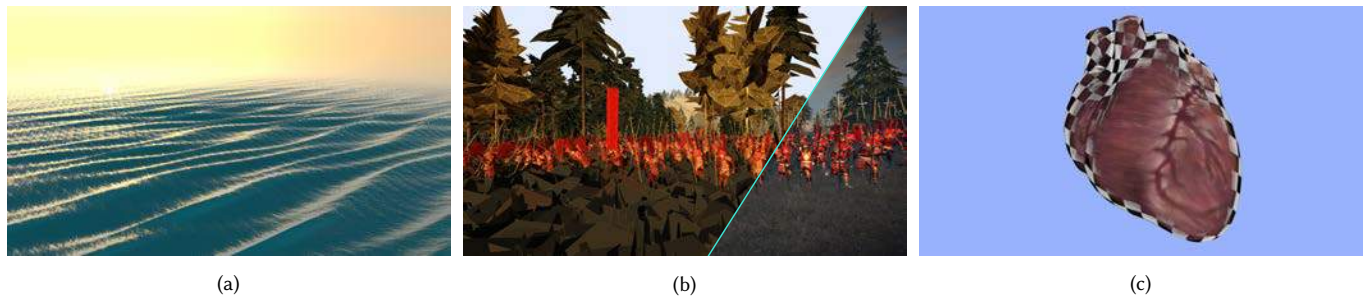


Fig. 1. Various scenes rendered by our software graphics pipeline in real-time on a GPU. (a) A smooth triangulation of the water surface in an animated ocean scene is achieved via a custom pipeline extension that allows the mesh topology to dynamically adapt to the underlying heightfield. (b) Scene geometry captured from video games like this still frame from *Total War: Shogun 2* is used to evaluate the performance of our approach on real-world triangle distributions. (c) Many techniques such as mipmapping rely on the ability to compute screen-space derivatives during fragment shading. Our pipeline architecture can support derivative estimation based on pixel quad shading, used here to render a textured model of a heart with trilinear filtering; lower mipmap levels are filled with a checkerboard pattern to visualize the effect. *Total War: Shogun 2* screenshot courtesy of *The Creative Assembly*; used with permission.

In this paper, we present a real-time graphics pipeline implemented entirely in software on a modern GPU. As opposed to previous work, our approach features a fully-concurrent, multi-stage, streaming design with dynamic load balancing, capable of operating efficiently within bounded memory. We address issues such as primitive order, vertex reuse, and screen-space derivatives of dependent variables, which are essential to real-world applications, but have largely been ignored by comparable work in the past. The power of a software approach lies in the ability to tailor the graphics pipeline to any given application. In exploration of this potential, we design and implement four novel pipeline modifications. Evaluation of the performance of our approach on more than 100 real-world scenes collected from video games shows rendering speeds within one order of magnitude of the hardware graphics pipeline as well as significant improvements over previous work, not only in terms of capabilities and performance, but also robustness.

CCS Concepts: • **Computing methodologies** → **Rasterization; Graphics processors**; *Massively parallel algorithms*;

Additional Key Words and Phrases: Software Rendering, GPU, Graphics Pipeline, Rasterization, CUDA

Authors' addresses: Michael Kenzel, michael.kenzel@icg.tugraz.at; Bernhard Kerbl, bernhard.kerbl@icg.tugraz.at; Dieter Schmalstieg, dieter.schmalstieg@icg.tugraz.at, Graz University of Technology, Institute of Computer Graphics and Vision, Inffeldgasse 16, Graz, 8010, Austria; Markus Steinberger, markus.steinberger@icg.tugraz.at, Graz University of Technology, Institute of Computer Graphics and Vision, Inffeldgasse 16, Graz, 8010, Austria, Max Planck Institute for Informatics, Saarland Informatics Campus Building E1 4, Saarbrücken, 66123, Germany.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART140 \$15.00

<https://doi.org/10.1145/3197517.3201374>

## ACM Reference Format:

Michael Kenzel, Bernhard Kerbl, Dieter Schmalstieg, and Markus Steinberger. 2018. A High-Performance Software Graphics Pipeline Architecture for the GPU. *ACM Trans. Graph.* 37, 4, Article 140 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201374>

## 1 INTRODUCTION

For a long time now, the hardware graphics pipeline has been the backbone of real-time rendering. However, while a hardware implementation can achieve high performance and power efficiency, flexibility is sacrificed. Driven by the need to support an ever growing spectrum of ever more sophisticated applications, the *graphics processing unit* (GPU) evolved as a tight compromise between flexibility and performance. The graphics pipeline on a modern GPU is implemented by special-purpose hardware on top of a large, freely-programmable, massively-parallel processor. More and more programmable stages have been added over the years. However, the overall structure of the pipeline and the underlying rendering algorithm have essentially remained unchanged for decades.

While evolution of the graphics pipeline proceeds slowly, GPU compute power continues to increase exponentially. In addition to the graphics pipeline, modern *application programming interfaces* (API) such as Vulkan [Khronos 2016b], OpenGL [Khronos 2016a], or Direct3D [Blythe 2006], as well as specialized interfaces like CUDA [NVIDIA 2016] and OpenCL [Stone et al. 2010] also allow the GPU to be operated in *compute mode*, which exposes the programmable cores of the GPU as a massively-parallel general-purpose co-processor. Although the hardware graphics pipeline remains at the core of real-time rendering, cutting-edge graphics applications increasingly rely on compute mode to implement major parts of sophisticated graphics algorithms that would not easily



# Slang: language mechanisms for extensible real-time shading systems

YONG HE, Carnegie Mellon University

KAYVON FATAHALIAN, Stanford University

TIM FOLEY, NVIDIA

Designers of real-time rendering engines must balance the conflicting goals of maintaining clear, extensible shading systems and achieving high rendering performance. In response, engine architects have established effective design patterns for authoring shading systems, and developed engine-specific code synthesis tools, ranging from preprocessor hacking to domain-specific shading languages, to productively implement these patterns. The problem is that proprietary tools add significant complexity to modern engines, lack advanced language features, and create additional challenges for learning and adoption. We argue that the advantages of engine-specific code generation tools can be achieved using the underlying GPU shading language directly, provided the shading language is extended with a small number of best-practice principles from modern, well-established programming languages. We identify that adding generics with interface constraints, associated types, and interface/structure extensions to existing C-like GPU shading languages enables real-time renderer developers to build shading systems that are extensible, maintainable, and execute efficiently on modern GPUs without the need for additional domain-specific tools. We embody these ideas in an extension of HLSL called Slang, and provide a reference design for a large, extensible shader library implemented using Slang's features. We rearchitect an open source renderer to use this library and Slang's compiler services, and demonstrate the resulting shading system is substantially simpler, easier to extend with new features, and achieves higher rendering performance than the original HLSL-based implementation.

CCS Concepts: • **Computing methodologies** → **Graphics systems and interfaces**;

Additional Key Words and Phrases: shading languages, real-time rendering

## ACM Reference Format:

Yong He, Kayvon Fatahalian, and Tim Foley. 2018. Slang: language mechanisms for extensible real-time shading systems. *ACM Trans. Graph.* 37, 4, Article 141 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201380>

## 1 INTRODUCTION

Designers of real-time rendering engines must balance the conflicting goals of facilitating developer productivity and achieving high rendering performance. Code maintainability and extensibility are key aspects of productivity, particularly since popular commercial engines such as Unreal [Epic Games 2015] or Unity [2017] feature large shader libraries used across many titles, each requiring different shading features. At the same time, to achieve high GPU

rendering performance, an engine must perform key optimizations such as statically specializing shader code to the rendering features in use, communicating shader parameter data between the CPU and GPU efficiently, and minimizing CPU overhead using the new parameter binding model offered by the modern Direct3D 12 and Vulkan graphics APIs.

To help navigate the tension between performance and maintainable/extensible code, engine architects have established effective design patterns for authoring shading systems, and developed code synthesis tools, ranging from preprocessor hacking, to metaprogramming, to engine-proprietary domain-specific languages (DSLs) [Tatarchuk and Tchou 2017], for implementing these patterns. For example, the idea of *shader components* [He et al. 2017] was recently presented as a pattern for achieving both high rendering performance and maintainable code structure when specializing shader code to coarse-grained features such as a surface material pattern or a tessellation effect. The idea of shader components is to drive both code specialization and CPU-GPU communication using the same granularity of decomposition, but the implementation of this idea was coupled to a custom DSL, which presents a barrier to adoption.

Beyond the coarse-grained specialization addressed by shader components, a modern shading system must be extensible to include new features and to assemble collections of shading effects into statically optimized GPU shaders. (Examples include: composing different types of lights, decoupling material patterns from reflectance models, adopting closed-form solutions and approximations for specific surface-light interactions.) AAA graphics engines implement proprietary code generation tools to aid with these tasks. Unfortunately, metaprogramming tools add complexity on top of the underlying shading language, and thus create additional challenges for learning and adoption. Engine-specific DSLs often lack advanced language features, and create the problem that shader code and learned skills do not transfer between engines.

In this paper we argue that the advantages of specialized DSLs (both the Spire language used to implement shader components [He et al. 2017] and other proprietary, engine-specific tools) can be achieved using the underlying GPU shading language directly, provided the shading language is extended with a small number of best-practice principles from modern, well-established programming languages. Our key contribution is to identify the necessary set of general-purpose modern programming language features that, when added to existing C-like GPU shading languages (GLSL/HLSL/Metal), can be used by real-time rendering engines to build shading systems that are extensible, maintainable, and execute efficiently on modern GPUs *without the need for additional layered DSLs*. Specifically, we:

- Propose the design of the *Slang shading language*, a variant of HLSL extended with the following general-purpose language features: generics with interface bounds, associated types, and interface/structure extensions. The choice of features is

Authors' addresses: Yong He, Carnegie Mellon University; Kayvon Fatahalian, Stanford University; Tim Foley, NVIDIA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART141 \$15.00  
<https://doi.org/10.1145/3197517.3201380>

# Learning Basketball Dribbling Skills Using Trajectory Optimization and Deep Reinforcement Learning

LIBIN LIU, DeepMotion Inc., USA

JESSICA HODGINS, Carnegie Mellon University, USA

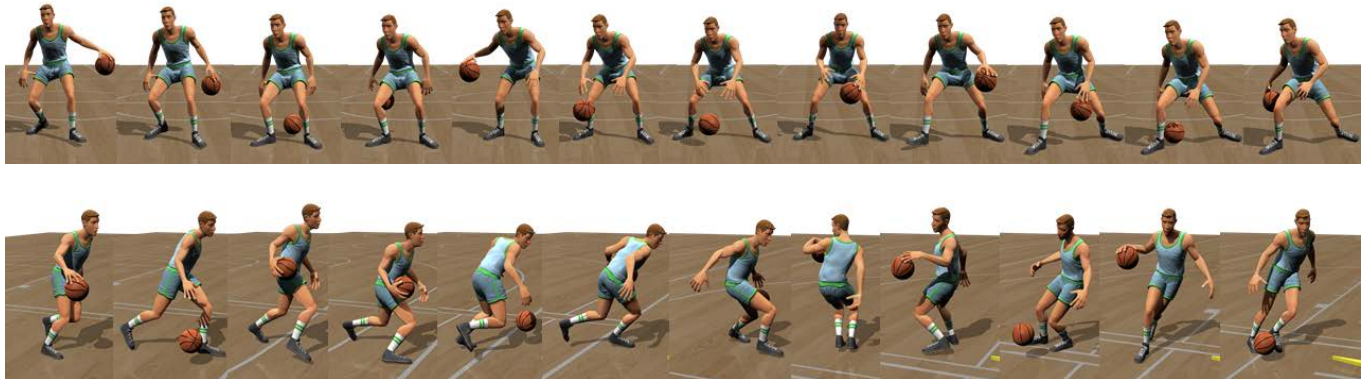


Fig. 1. Real-time simulation of the learned basketball skills. Top: A player dribbles a ball behind the back and between the legs. Bottom: A player performs crossover moves.

Basketball is one of the world's most popular sports because of the agility and speed demonstrated by the players. This agility and speed makes designing controllers to realize robust control of basketball skills a challenge for physics-based character animation. The highly dynamic behaviors and precise manipulation of the ball that occur in the game are difficult to reproduce for simulated players. In this paper, we present an approach for learning robust basketball dribbling controllers from motion capture data. Our system decouples a basketball controller into locomotion control and arm control components and learns each component separately. To achieve robust control of the ball, we develop an efficient pipeline based on trajectory optimization and deep reinforcement learning and learn non-linear arm control policies. We also present a technique for learning skills and the transition between skills simultaneously. Our system is capable of learning robust controllers for various basketball dribbling skills, such as dribbling between the legs and crossover moves. The resulting control graphs enable a simulated player to perform transitions between these skills and respond to user interaction.

**CCS Concepts:** • **Computing methodologies** → **Physical simulation;** *Neural networks; Motion capture;*

Authors' addresses: Libin Liu, DeepMotion Inc. 3 Twin Dolphin Dr. Suite 295, Redwood City, CA, 94065, USA, libin@deepmotion.com; Jessica Hodgins, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA, 15213, USA, jkh@cs.cmu.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART142 \$15.00  
<https://doi.org/10.1145/3197517.3201315>

**Additional Key Words and Phrases:** physics-based characters, human simulation, motion control, basketball, deep deterministic policy gradient

## ACM Reference Format:

Libin Liu and Jessica Hodgins. 2018. Learning Basketball Dribbling Skills Using Trajectory Optimization and Deep Reinforcement Learning. *ACM Trans. Graph.* 37, 4, Article 142 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201315>

## 1 INTRODUCTION

Basketball is one of the world's most popular sports. Animating basketball sports with motion capture data is challenging because the precise coupling between the movement of the ball and the motion of the player can be easily damaged during kinematic operations such as blending and deformation. Although the state-of-the-art basketball video games, such as the NBA 2K series and the NBA Live series, have demonstrated very natural motion with motion capture data and have provided a lot of enjoyment to fans, they still contain artifacts. For example, the ball may move in a physically implausible trajectory or appear to be stuck to the player's hand. Designing physics-based controllers for basketball skills has the potential to create high-quality, physically realistic basketball animation. However, designing such controllers is challenging because of the details of the physical interactions with the ball and the ground, the agility and grace required to perform the required tasks, and the coupling between the locomotion control and the manipulation of the fast-moving basketball.

In this paper, we describe a method for learning robust basketball dribbling controllers for simulated players from motion capture data. Our method treats a basketball controller as a combination of locomotion control and arm control components. The system learns

# DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills

XUE BIN PENG, University of California, Berkeley  
PIETER ABBEEL, University of California, Berkeley  
SERGEY LEVINE, University of California, Berkeley  
MICHIEL VAN DE PANNE, University of British Columbia

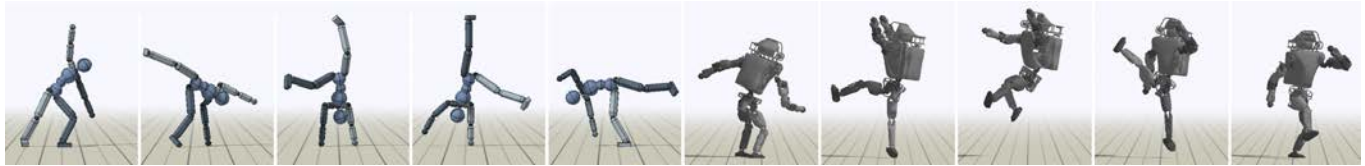


Fig. 1. Highly dynamic skills learned by imitating reference motion capture clips using our method, executed by physically simulated characters. **Left:** Humanoid character performing a cartwheel. **Right:** Simulated Atlas robot performing a spinkick.

A longstanding goal in character animation is to combine data-driven specification of behavior with a system that can execute a similar behavior in a physical simulation, thus enabling realistic responses to perturbations and environmental variation. We show that well-known reinforcement learning (RL) methods can be adapted to learn robust control policies capable of imitating a broad range of example motion clips, while also learning complex recoveries, adapting to changes in morphology, and accomplishing user-specified goals. Our method handles keyframed motions, highly-dynamic actions such as motion-captured flips and spins, and retargeted motions. By combining a motion-imitation objective with a task objective, we can train characters that react intelligently in interactive settings, e.g., by walking in a desired direction or throwing a ball at a user-specified target. This approach thus combines the convenience and motion quality of using motion clips to define the desired style and appearance, with the flexibility and generality afforded by RL methods and physics-based animation. We further explore a number of methods for integrating multiple clips into the learning process to develop multi-skilled agents capable of performing a rich repertoire of diverse skills. We demonstrate results using multiple characters (human, Atlas robot, bipedal dinosaur, dragon) and a large variety of skills, including locomotion, acrobatics, and martial arts.

CCS Concepts: • **Computing methodologies** → **Animation**; *Physical simulation*; *Control methods*; *Reinforcement learning*;

Additional Key Words and Phrases: physics-based character animation, motion control, reinforcement learning

## ACM Reference Format:

Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. 2018. DeepMimic: Example-Guided Deep Reinforcement Learning of Physics-Based Character Skills. *ACM Trans. Graph.* 37, 4, Article 143 (August 2018), 18 pages. <https://doi.org/10.1145/3197517.3201311>

Authors' addresses: Xue Bin Peng, University of California, Berkeley; Pieter Abbeel, University of California, Berkeley; Sergey Levine, University of California, Berkeley; Michiel van de Panne, University of British Columbia.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2018/8-ART143 \$15.00

<https://doi.org/10.1145/3197517.3201311>

## 1 INTRODUCTION

Physics-based simulation of passive phenomena, such as cloth and fluids, has become nearly ubiquitous in industry. However, the adoption of physically simulated characters has been more modest. Modeling the motion of humans and animals remains a challenging problem, and currently, few methods exist that can simulate the diversity of behaviors exhibited in the real world. Among the enduring challenges in this domain are generalization and directability. Methods that rely on manually designed controllers have produced compelling results, but their ability to generalize to new skills and new situations is limited by the availability of human insight. Though humans are adept at performing a wide range of skills themselves, it can be difficult to articulate the internal strategies that underly such proficiency, and more challenging still to encode them into a controller. Directability is another obstacle that has impeded the adoption of simulated characters. Authoring motions for simulated characters remains notoriously difficult, and current interfaces still cannot provide users with an effective means of eliciting the desired behaviours from simulated characters.

Reinforcement learning (RL) provides a promising approach for motion synthesis, whereby an agent learns to perform various skills through trial-and-error, thus reducing the need for human insight. While deep reinforcement learning has been demonstrated to produce a range of complex behaviors in prior work [Duan et al. 2016; Heess et al. 2016; Schulman et al. 2015b], the quality of the generated motions has thus far lagged well behind state-of-the-art kinematic methods or manually designed controllers. In particular, controllers trained with deep RL exhibit severe (and sometimes humorous) artifacts, such as extraneous upper body motion, peculiar gaits, and unrealistic posture [Heess et al. 2017].<sup>1</sup> A natural direction to improve the quality of learned controllers is to incorporate motion capture or hand-authored animation data. In prior work, such systems have typically been designed by layering a physics-based tracking controller on top of a kinematic animation system [Da Silva et al. 2008; Lee et al. 2010a]. This type of approach is challenging because the kinematic animation system must produce reference motions that

<sup>1</sup>See, for example, [https://youtu.be/hx\\_bgoTF7bs](https://youtu.be/hx_bgoTF7bs)



# Learning Symmetric and Low-Energy Locomotion

WENHAO YU, Georgia Institute of Technology

GREG TURK, Georgia Institute of Technology

C.KAREN LIU, Georgia Institute of Technology

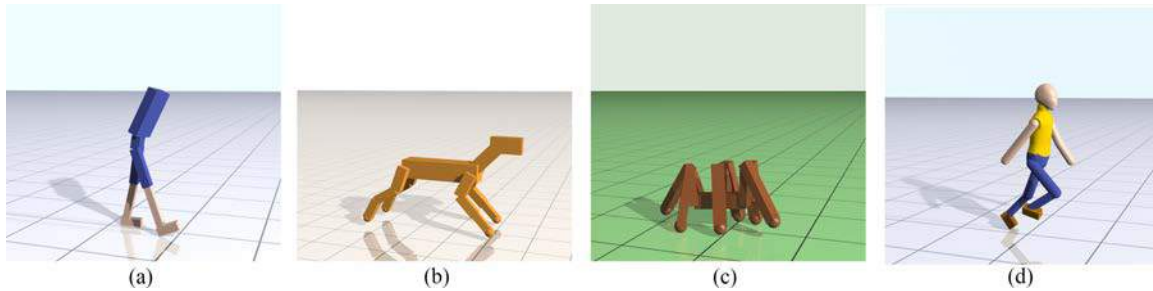


Fig. 1. Locomotion Controller trained for different creatures. (a) Biped walking. (b) Quadruped galloping. (c) Hexapod Walking. (d) Humanoid running.

Learning locomotion skills is a challenging problem. To generate realistic and smooth locomotion, existing methods use motion capture, finite state machines or morphology-specific knowledge to guide the motion generation algorithms. Deep reinforcement learning (DRL) is a promising approach for the automatic creation of locomotion control. Indeed, a standard benchmark for DRL is to automatically create a running controller for a biped character from a simple reward function [Duan et al. 2016]. Although several different DRL algorithms can successfully create a running controller, the resulting motions usually look nothing like a real runner. This paper takes a minimalist learning approach to the locomotion problem, without the use of motion examples, finite state machines, or morphology-specific knowledge. We introduce two modifications to the DRL approach that, when used together, produce locomotion behaviors that are symmetric, low-energy, and much closer to that of a real person. First, we introduce a new term to the loss function (not the reward function) that encourages symmetric actions. Second, we introduce a new curriculum learning method that provides modulated physical assistance to help the character with left/right balance and forward movement. The algorithm automatically computes appropriate assistance to the character and gradually relaxes this assistance, so that eventually the character learns to move entirely without help. Because our method does not make use of motion capture data, it can be applied to a variety of character morphologies. We demonstrate locomotion controllers for the lower half of a biped, a full humanoid, a quadruped, and a hexapod. Our results show that learned policies are able to produce symmetric, low-energy gaits. In addition, speed-appropriate gait patterns emerge without any guidance from motion examples or contact planning.

Authors' addresses: Wenhao Yu, School of Interactive Computing, Georgia Institute of Technology, wyu68@gatech.edu; Greg Turk, School of Interactive Computing, Georgia Institute of Technology, turk@cc.gatech.edu; C.Karen Liu, School of Interactive Computing, Georgia Institute of Technology, karenliu@cc.gatech.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART144 \$15.00

<https://doi.org/10.1145/3197517.3201397>

CCS Concepts: • **Computing methodologies** → **Animation; Reinforcement Learning**;

Additional Key Words and Phrases: locomotion, reinforcement learning, curriculum learning

## ACM Reference Format:

Wenhao Yu, Greg Turk, and C.Karen Liu. 2018. Learning Symmetric and Low-Energy Locomotion. *ACM Trans. Graph.* 37, 4, Article 144 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201397>

## 1 INTRODUCTION

Creating an animated character that can walk is a fascinating challenge for graphics researchers and animators. Knowledge from biomechanics, physics, robotics, and animation give us ideas for how to coordinate the virtual muscles of a character's body to move it forward while maintaining balance and style. Whether physical or digital, the creators of these characters apply physics principles, borrow ideas from domain experts, use motion data, and undertake arduous trial-and-error to craft lifelike movements that mimic real-world animals and humans. While these characters can be engineered to exhibit locomotion behaviors, a more intriguing question is whether the characters can learn locomotion behaviors on their own, the way a human toddler can.

The recent disruptive development in Deep Reinforcement Learning (DRL) suggests that the answer is yes. Researchers indeed showed that artificial agents can learn some form of locomotion using advanced policy learning methods with a large amount of computation. Even though such agents are able to move from point A to point B without falling, the resulting motion usually exhibits jerky, high-frequency movements. The motion artifacts can be mitigated by introducing motion examples or special objectives in the reward function, but these remedies are somewhat unsatisfying as they sacrifice generality of locomotion principles and return partway to heavily engineered solutions.

This paper takes a minimalist approach to the problem of learning locomotion. Our hypothesis is that natural locomotion will emerge from simple and well-known principles in biomechanics, without

# Mode-Adaptive Neural Networks for Quadruped Motion Control

HE ZHANG<sup>†</sup>, University of Edinburgh  
SEBASTIAN STARKE<sup>†</sup>, University of Edinburgh  
TAKU KOMURA, University of Edinburgh  
JUN SAITO, Adobe Research

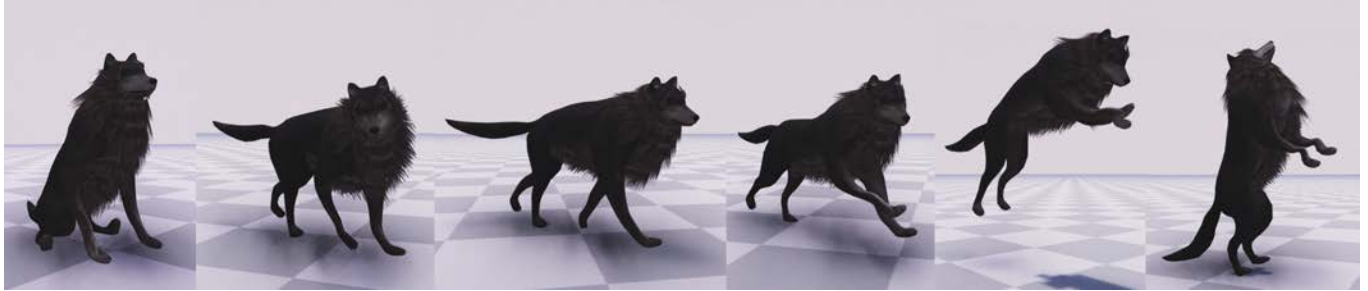


Fig. 1. A selection of results using our method for quadruped animation. We show some different modes for sitting, turning trot, pace, canter, jumping and standing from left to right. The locomotion gaits are not labeled individually, but naturally produced by the movement velocity control.

Quadruped motion includes a wide variation of gaits such as walk, pace, trot and canter, and actions such as jumping, sitting, turning and idling. Applying existing data-driven character control frameworks to such data requires a significant amount of data preprocessing such as motion labeling and alignment. In this paper, we propose a novel neural network architecture called Mode-Adaptive Neural Networks for controlling quadruped characters. The system is composed of the motion prediction network and the gating network. At each frame, the motion prediction network computes the character state in the current frame given the state in the previous frame and the user-provided control signals. The gating network dynamically updates the weights of the motion prediction network by selecting and blending what we call the expert weights, each of which specializes in a particular movement. Due to the increased flexibility, the system can learn consistent expert weights across a wide range of non-periodic/periodic actions, from unstructured motion capture data, in an end-to-end fashion. In addition, the users are released from performing complex labeling of phases in different gaits. We show that this architecture is suitable for encoding the multi-modality of quadruped locomotion and synthesizing responsive motion in real-time.

CCS Concepts: • **Computing methodologies** → **Motion capture**; *Neural networks*;

<sup>†</sup> He Zhang and Sebastian Starke are joint first authors.

Authors' addresses: He Zhang<sup>†</sup>, University of Edinburgh, s1676928@sms.ed.ac.uk; Sebastian Starke<sup>†</sup>, University of Edinburgh, sebastian.starke@ed.ac.uk; Taku Komura, University of Edinburgh, tkomura@ed.ac.uk; Jun Saito, Adobe Research, jsaito@adobe.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.  
0730-0301/2018/8-ART145 \$15.00  
<https://doi.org/10.1145/3197517.3201366>

Additional Key Words and Phrases: neural networks, locomotion, human motion, character animation, character control, deep learning

## ACM Reference Format:

He Zhang<sup>†</sup>, Sebastian Starke<sup>†</sup>, Taku Komura, and Jun Saito. 2018. Mode-Adaptive Neural Networks for Quadruped Motion Control. *ACM Trans. Graph.* 37, 4, Article 145 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201366>

## 1 INTRODUCTION

Quadruped animation is one of the unsolved key problems in computer animation. It has particular relevance for applications in computer games and films, and also presents a challenging topic in robotics. When animating quadrupeds, animators must go through special training to design a wide range of complex movements. This complexity is inherently from the multi-modality of quadruped motions. For example, there are a number of locomotion modes<sup>1</sup> including walk, pace, trot, canter, and gallop, where the movements and the phases of the torso and limbs vary in a complex manner (see Fig. 2).

To the best of our knowledge, there is no prior work on systematically constructing quadruped motion controllers in a data-driven fashion. This difficulty stems from the complexity mentioned above, along with the issue that quadruped animals cannot be directed like humans for a controlled data acquisition. As a result, the captured data is often less structured with a wide range of random actions performed one after another. When designing character controllers using such data, engineers need to manually/semi-automatically extract gait cycles and transitions from the data, stitch them together, and tune parameters of motion trees and finite state machines.

<sup>1</sup>Locomotion modes is the term to represent the variation in movements the animals use to propel, here which is defined based on the binary footfall patterns along the timeline that the end effectors contact the ground. "Locomotion modes" and "gait types" are interchangeably used in this article.

# Physical Simulation of Environmentally Induced Thin Shell Deformation

HSIAO-YU CHEN, Institute for Computational Engineering and Sciences, University of Texas at Austin

ARNAV SASTRY, Department of Computer Science, University of Texas at Austin

WIM M. VAN REES, Department of Mechanical Engineering, Massachusetts Institute of Technology

ETIENNE VOUGA, Department of Computer Science, University of Texas at Austin

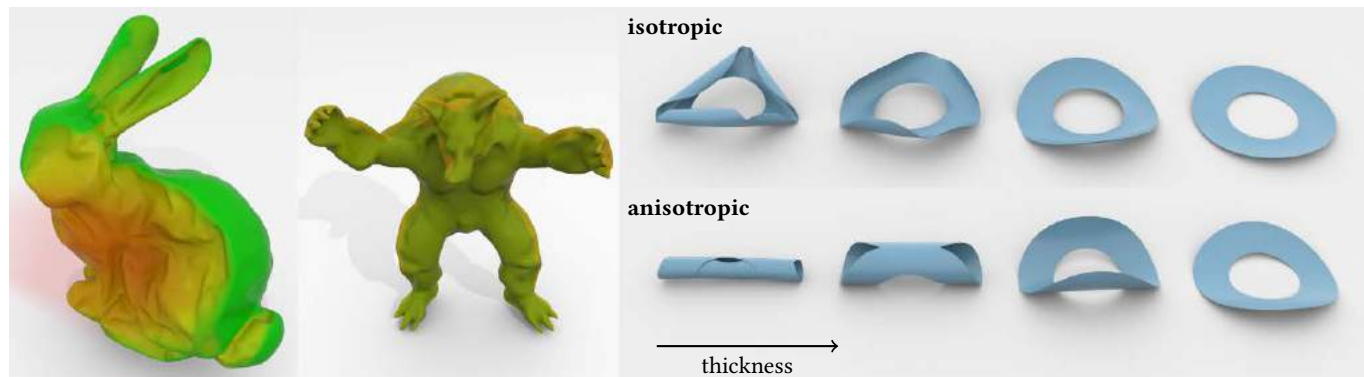


Fig. 1. The geometry of thin materials often changes dramatically in response to heat and moisture. A plastic bunny and armadillo shrink in response to being heated (left). Colors on the surfaces indicate temperature (redder is hotter). Moist paper annuli curl up as they dry, and the amount and shape of curling is based on the paper's thickness and whether or not the paper's fibers are randomly or consistently oriented (right).

We present a physically accurate low-order elastic shell model that incorporates active material response to dynamically changing stimuli such as heat, moisture, and growth. Our continuous formulation of the geometrically non-linear elastic energy derives from the principles of differential geometry, and as such naturally incorporates shell thickness, non-zero rest curvature, and physical material properties. By modeling the environmental stimulus as local, dynamic changes in the rest metric of the material, we are able to solve for the corresponding shape changes by integrating the equations of motions given this non-Euclidean rest state. We present models for differential growth and shrinking due to moisture and temperature gradients along and across the surface, and incorporate anisotropic growth by defining an intrinsic machine direction within the material. Comparisons with experiments and volumetric finite elements show that our simulations achieve excellent qualitative and quantitative agreement. By combining the reduced-order shell theory with appropriate physical models, our approach

accurately captures all the physical phenomena while avoiding expensive volumetric discretization of the shell volume.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: differential growth, thin shell

## ACM Reference Format:

Hsiao-yu Chen, Arnav Sastry, Wim M. van Rees, and Etienne Vouga. 2018. Physical Simulation of Environmentally Induced Thin Shell Deformation. *ACM Trans. Graph.* 37, 4, Article 146 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201395>

## 1 INTRODUCTION

Consider the wrinkling and curling of a drying leaf. The drying process corresponds to water evaporating from the internal cells, so that the tissue contracts in volume. This process typically happens differentially, due to having one side exposed to the sun, or having boundaries farther away from the veins than interior, leading to non-uniform curvature and curling of the leaf [Jeong et al. 2013; Xiao and Chen 2011]. Related phenomena are the swelling and wrinkling of paper when exposed to water, such as when coffee is spilled on a notebook page, or the wrinkling and shrinking of plastic when heated. Although the physical mechanisms in these examples are different (shrinking of cells vs swelling of fibers vs contraction of polymers) the *effect* on the objects is the same: the intrinsic geometry of the thin objects change over time in response to dynamic changes in the environment.

In addition to growth, wrinkling, and swelling, other phenomena involving the same physics include burning of thin objects, especially those made of curved or composite materials; wilting of

Authors' addresses: Hsiao-yu Chen, Institute for Computational Engineering and Sciences, University of Texas at Austin, [hsiaoyu@ices.utexas.edu](mailto:hsiaoyu@ices.utexas.edu); Arnav Sastry, Department of Computer Science, University of Texas at Austin, [arnavsastry@gmail.com](mailto:arnavsastry@gmail.com); Wim M. van Rees, Department of Mechanical Engineering, Massachusetts Institute of Technology, [wvanrees@mit.edu](mailto:wvanrees@mit.edu); Etienne Vouga, Department of Computer Science, University of Texas at Austin, [evouga@cs.utexas.edu](mailto:evouga@cs.utexas.edu).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART146 \$15.00

<https://doi.org/10.1145/3197517.3201395>



# A Material Point Method for Thin Shells with Frictional Contact

QI GUO, University of California, Los Angeles

XUCHEN HAN, University of California, Los Angeles

CHUYUAN FU, University of California, Los Angeles

THEODORE GAST, University of California, Los Angeles

RASMUS TAMSTORF, Walt Disney Animation Studios

JOSEPH TERAN, University of California, Los Angeles



Fig. 1. **Shell Montage.** Upper left : simulation of shells coupled with granular materials. Center left : a walk cycle benchmark for clothing simulation. Bottom right : a T-shirt twisted to induce many self-collisions. Center : the effect of increasing bending stiffness (from left to right) for six collapsing elastic cylinders.

Authors' addresses: Qi Guo, University of California, Los Angeles, qiguo@ucla.edu; Xuchen Han, University of California, Los Angeles, xhan0619@ucla.edu; Chuyuan Fu, University of California, Los Angeles, fuchuyuan@math.ucla.edu; Theodore Gast, University of California, Los Angeles, tfg@math.ucla.edu; Rasmus Tamstorf, Walt Disney Animation Studios, rasmus.tamstorf@disneyanimation.com; Joseph Teran, University of California, Los Angeles, jteran@math.ucla.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://www.acm.org/permissions).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART147 \$15.00

<https://doi.org/10.1145/3197517.3201346>

We present a novel method for simulation of thin shells with frictional contact using a combination of the Material Point Method (MPM) and subdivision finite elements. The shell kinematics are assumed to follow a continuum shell model which is decomposed into a Kirchhoff-Love motion that rotates the mid-surface normals followed by shearing and compression/extension of the material along the mid-surface normal. We use this decomposition to design an elastoplastic constitutive model to resolve frictional contact by decoupling resistance to contact and shearing from the bending resistance components of stress. We show that by resolving frictional contact with a continuum approach, our hybrid Lagrangian/Eulerian approach is capable of simulating challenging shell contact scenarios with hundreds of thousands to millions of degrees of freedom. Without the need for collision detection or resolution, our method runs in a few minutes per frame in these high resolution examples. Furthermore we show that our technique naturally couples with other traditional MPM methods for simulating granular and related materials.

# Mechanical Characterization of Structured Sheet Materials

CHRISTIAN SCHUMACHER, Disney Research and ETH Zurich

STEVE MARSCHNER, Cornell University

MARKUS GROSS, Disney Research and ETH Zurich

BERNHARD THOMASZEWSKI, Université de Montréal

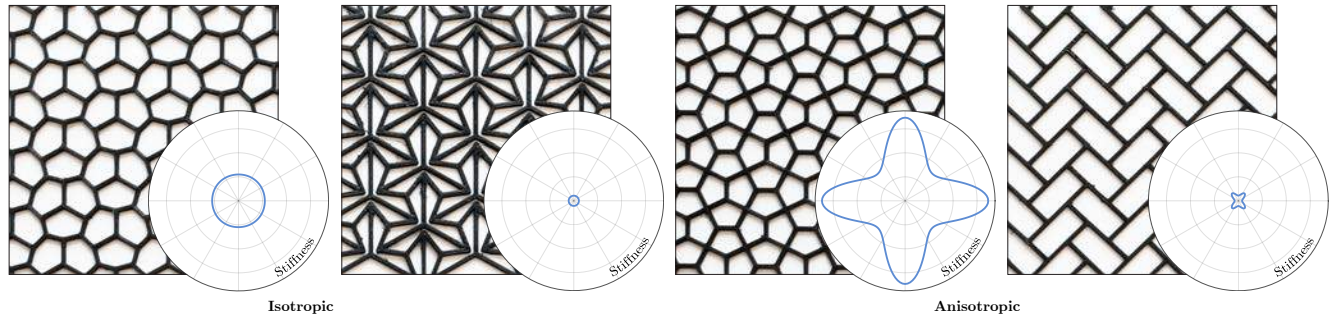


Fig. 1. 3D-printed tilings combine aesthetically-pleasing geometry with complex mechanical behavior characterized by anisotropy, nonlinearity, and large variations in stiffness among different patterns. Using numerical homogenization, our method summarizes the macro-mechanical behavior of these tilings in terms of direction-dependent Young's modulus (shown above), Poisson's ratio, and bending stiffness as well as a simple measure of nonlinearity.

We propose a comprehensive approach to characterizing the mechanical properties of structured sheet materials, i.e., planar rod networks whose mechanics and aesthetics are inextricably linked. We establish a connection between the complex mesoscopic deformation behavior of such structures and their macroscopic elastic properties through numerical homogenization. Our approach leverages 3D Kirchhoff rod simulation in order to capture nonlinear effects for both in-plane and bending deformations. We apply our method to different families of structures based on isohedral tilings—a simple yet extensive and aesthetically interesting group of space-filling patterns. We show that these tilings admit a wide range of material properties, and our homogenization approach allows us to create concise and intuitive descriptions of a material's direction-dependent macromechanical behavior that are easy to communicate even to non-experts. We perform this characterization for an extensive set of structures and organize these data in a material browser to enable efficient forward exploration of the aesthetic-mechanical space of structured sheet materials. We also propose an inverse design method to automatically find structure parameters that best approximate a user-specified target behavior.

CCS Concepts: • **Applied computing** → **Computer-aided design**;

Additional Key Words and Phrases: Mechanical Characterization, Space-Filling Tiling

Authors' addresses: Christian Schumacher, Disney Research and ETH Zurich; Steve Marschner, Cornell University; Markus Gross, Disney Research and ETH Zurich; Bernhard Thomaszewski, Université de Montréal.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART148 \$15.00

<https://doi.org/10.1145/3197517.3201278>

## ACM Reference Format:

Christian Schumacher, Steve Marschner, Markus Gross, and Bernhard Thomaszewski. 2018. Mechanical Characterization of Structured Sheet Materials. *ACM Trans. Graph.* 37, 4, Article 148 (August 2018), 15 pages. <https://doi.org/10.1145/3197517.3201278>

## 1 INTRODUCTION

Simple shapes can combine into structures of fascinating regularity and complexity. Thanks to their aesthetic appeal and diversity, such *tessellations* are used as a universal way of decorating buildings, garments, and everyday things across many cultures. But the diversity created by these patterns is not limited to appearance. Just as aesthetic complexity can be created from simple shapes, so can mechanical complexity. Indeed, physical representations of these patterns—which we call *structured sheet materials*—exhibit a wide range of macromechanical properties mediated only by geometry and topology. Unlike appearance, however, the mechanical properties of structured sheet materials are often far from obvious. Exploring and understanding the connection between the geometry and mechanics of structured sheet materials is the main objective of this work.

The graphics community has recently started to explore the design of flexible materials whose microstructure can be controlled in order to achieve desired macroscopic deformation behavior [Bickel et al. 2010; Martínez et al. 2016; Panetta et al. 2015; Schumacher et al. 2015]. While previous work has so far focused on volumetric materials, our structured sheets are planar networks of thin elastic rods that can stretch and *bend*. We focus on isohedral polygonal tilings, a particular class of periodic patterns in which all tiles are congruent to a single polygon. Polygonal tilings have been studied intensively in mathematics [Grünbaum and Shephard 1986] and graphics [Kaplan 2009]. We show that this geometrically rich space

# Animating Fluid Sediment Mixture in Particle-Laden Flows

MING GAO, University of Wisconsin, Madison, USA

ANDRE PRADHANA, University of Pennsylvania, USA

XUCHEN HAN, University of California, Los Angeles, USA

QI GUO, University of California, Los Angeles, USA

GRANT KOT, Phosphorus, USA

EFTYCHIOS SIFAKIS, University of Wisconsin, Madison, USA

CHENFANFU JIANG, University of Pennsylvania, USA



Fig. 1. **Sediment transport:** Our method can animate intricate two-way coupled particle-laden flows such as sediment transport in liquid.

In this paper, we present a mixed explicit and semi-implicit Material Point Method for simulating particle-laden flows. We develop a Multigrid Preconditioned fluid solver for the Locally Averaged Navier Stokes equation. This is discretized purely on a semi-staggered standard MPM grid. Sedimentation is modeled with the Drucker-Prager elastoplasticity flow rule, enhanced by a novel particle density estimation method for converting particles between representations of either continuum or discrete points. Fluid and sediment are two-way coupled through a momentum exchange force that can be easily resolved with two MPM background grids. We present various results to demonstrate the efficacy of our method.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Additional Key Words and Phrases: Material Point Method (MPM), particle-fluid interaction, multiphase, sedimentation, sediment transport

Authors' addresses: Ming Gao, University of Wisconsin, Madison, USA; Andre Pradhana, University of Pennsylvania, Philadelphia, USA; Xuchen Han, University of California, Los Angeles, USA; Qi Guo, University of California, Los Angeles, USA; Grant Kot, Phosphorus, USA; Eftychios Sifakis, University of Wisconsin, Madison, USA; Chenfanfu Jiang, University of Pennsylvania, Philadelphia, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART149 \$15.00

<https://doi.org/10.1145/3197517.3201309>

## ACM Reference Format:

Ming Gao, Andre Pradhana, Xuchen Han, Qi Guo, Grant Kot, Eftychios Sifakis, and Chenfanfu Jiang. 2018. Animating Fluid Sediment Mixture in Particle-Laden Flows. *ACM Trans. Graph.* 37, 4, Article 149 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201309>

## 1 INTRODUCTION

Recently, multi-phase multi-material simulations are increasingly gaining attention from computer graphics researchers. Simulating various phases or materials in a unified framework is particularly favored. Existing work includes coupled Lagrangian particle simulation with Position Based Dynamics (PBD) [Macklin et al. 2014], water-gas mixtures [Nielsen and Østerby 2013] with an Eulerian method, solid-fluid phase-change [Stomakhin et al. 2014] and porous granular media [Pradhana-Tampubolon et al. 2017] with Material Point Method (MPM), as well as interactive solids and fluids based on the mixture model with Smoothed Particle Hydrodynamics (SPH) [Yan et al. 2016].

Most of the existing approaches are based on *continuum* mixture theory [Manninen et al. 1996]. The continuum assumption for each material phase is essential for simulations of macroscopic porous media (e.g., landslides and liquid blending). However, it may fail to capture the correct behavior of particle-laden flows where the solid phase is on a relatively small scale. Note that particle-laden sediment flow is ubiquitous in natural systems. Typical examples include sediment transport, sedimentation, volcano eruption, dune



# A Moving Least Squares Material Point Method with Displacement Discontinuity and Two-Way Rigid Body Coupling

YUANMING HU<sup>†</sup>, MIT CSAIL

YU FANG<sup>†</sup>, Tsinghua University

ZIHENG GE<sup>†</sup>, University of Science and Technology of China

ZIYIN QU, University of Pennsylvania

YIXIN ZHU<sup>†</sup>, University of California, Los Angeles

ANDRE PRADHANA, University of Pennsylvania

CHENFANFU JIANG, University of Pennsylvania

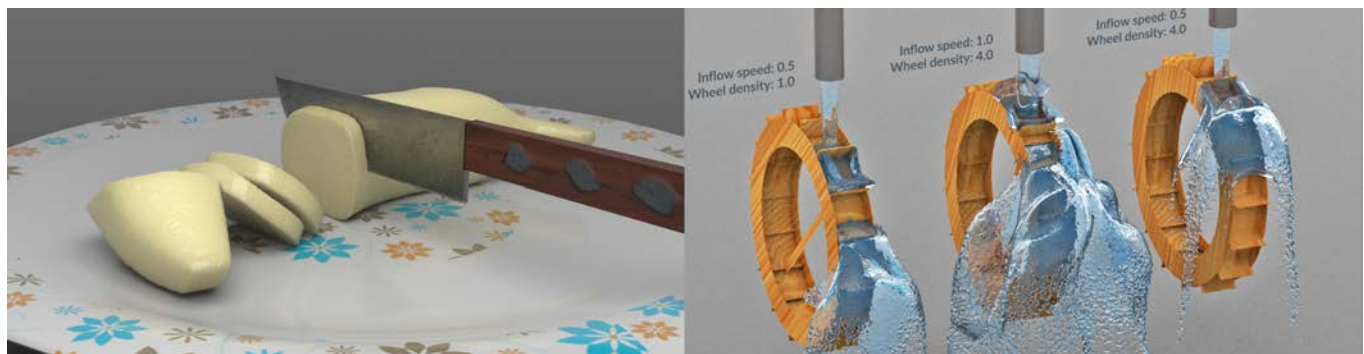


Fig. 1. Our method allows MPM to handle world space material cutting, complex thin boundaries and natural two-way rigid body coupling.

In this paper, we introduce the Moving Least Squares Material Point Method (MLS-MPM). MLS-MPM naturally leads to the formulation of Affine Particle-In-Cell (APIC) [Jiang et al. 2015] and Polynomial Particle-In-Cell [Fu et al. 2017] in a way that is consistent with a Galerkin-style weak form discretization of the governing equations. Additionally, it enables a new stress divergence discretization that effortlessly allows all MPM simulations to run two times faster than before. We also develop a Compatible Particle-In-Cell (CPIC) algorithm on top of MLS-MPM. Utilizing a colored distance field representation and a novel compatibility condition for particles and grid nodes, our framework enables the simulation of various new phenomena that are not previously supported by MPM, including material cutting, dynamic open boundaries, and two-way coupling with rigid bodies. MLS-MPM with CPIC is easy to implement and friendly to performance optimization.

CCS Concepts: • **Computing methodologies** → **Physical simulation**;

Authors' addresses: Yuanming Hu<sup>†</sup>, MIT CSAIL, yuanming@mit.edu; Yu Fang<sup>†</sup>, Tsinghua University, squarefk@gmail.com; Ziheng Ge<sup>†</sup>, University of Science and Technology of China, gzh1057@mail.ustc.edu.cn; Ziyin Qu, University of Pennsylvania, ziyinq@seas.upenn.edu; Yixin Zhu<sup>†</sup>, University of California, Los Angeles, yixin.zhu@ucla.edu; Andre Pradhana, University of Pennsylvania, apradh@seas.upenn.edu; Chenfanfu Jiang, University of Pennsylvania, cffjiang@seas.upenn.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](https://www.acm.org/permissions).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART150 \$15.00

<https://doi.org/10.1145/3197517.3201293>

Additional Key Words and Phrases: Material Point Method (MPM), moving least squares, cutting, discontinuity, distance field, rigid coupling

## ACM Reference Format:

Yuanming Hu<sup>†</sup>, Yu Fang<sup>†</sup>, Ziheng Ge<sup>†</sup>, Ziyin Qu, Yixin Zhu<sup>†</sup>, Andre Pradhana, and Chenfanfu Jiang. 2018. A Moving Least Squares Material Point Method with Displacement Discontinuity and Two-Way Rigid Body Coupling. *ACM Trans. Graph.* 37, 4, Article 150 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201293>

## 1 INTRODUCTION

Since the pioneering work of Terzopoulos et al. [1988], simulating topologically changing materials has been a popular research topic in graphics. Among various topics, fracture and cutting of deformable objects are most intensively explored. The goal of breaking mesh connectivity has led to techniques such as local remeshing [O'Brien et al. 2002; O'Brien and Hodgins 1999], the Virtual Node Algorithm (VNA) [Hegemann et al. 2013; Molino et al. 2005; Wang et al. 2014] and the eXtended Finite Element Method (XFEM) [Koschier et al. 2017]. Maintaining remeshing quality efficiently and robustly can be complicated. While VNA and XFEM reduce some difficulty, they impose additional challenges like floating point arithmetic in degenerate scenarios and self-collision on embedded surfaces.

Compared to mesh-based approaches, meshless animation of solid topology change was shown to be promising by Pauly et al. [2005]. More recently, the Material Point Method (MPM) [Sulsky et al. 1995]

<sup>†</sup> Y. Hu, Y. Fang, Z. Ge and Y. Zhu were visiting students at the University of Pennsylvania during this work.

# Predictive and Generative Neural Networks for Object Functionality

RUIZHEN HU, Shenzhen University  
 ZIHAO YAN, Shenzhen University  
 JINGWEN ZHANG, Shenzhen University  
 OLIVER VAN KAICK, Carleton University  
 ARIEL SHAMIR, The Interdisciplinary Center  
 HAO ZHANG, Simon Fraser University  
 HUI HUANG\*, Shenzhen University

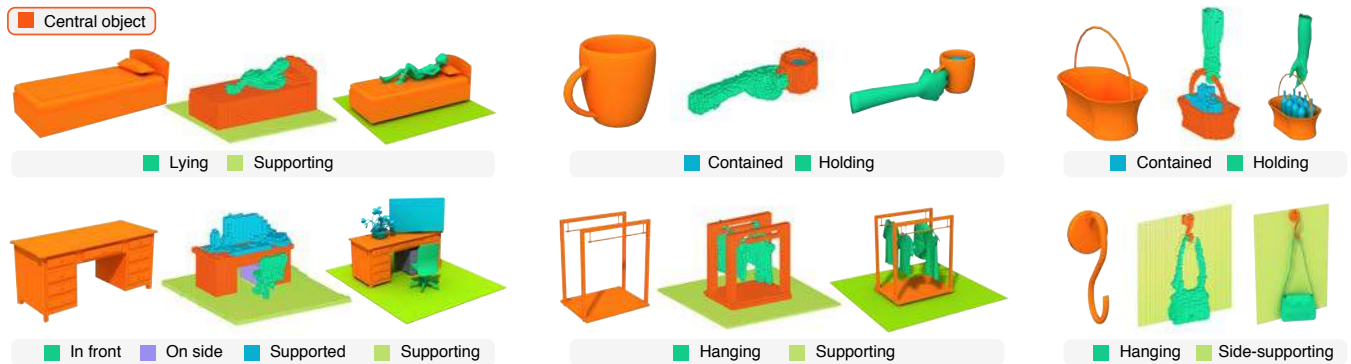


Fig. 1. Given an object in isolation (left of each example), our generative network synthesizes scenes that demonstrate the functionality of the object in terms of interactions with surrounding objects (middle). Note the different types of functionalities appearing in the scenes generated by the network, involving interactions such as support, containment, and grasping. The scene is refined by replacing voxelized objects with higher resolution models (right).

Humans can predict the functionality of an object even without any surroundings, since their knowledge and experience would allow them to “hallucinate” the interaction or usage scenarios involving the object. We develop *predictive* and *generative* deep convolutional neural networks to replicate this feat. Specifically, our work focuses on functionalities of man-made 3D objects characterized by human-object or object-object interactions. Our networks are trained on a database of scene contexts, called *interaction contexts*, each consisting of a central object and one or more surrounding objects, that represent object functionalities. Given a 3D object *in isolation*, our *functional similarity network* (fSIM-NET), a variation of the triplet network, is trained to predict the functionality of the object by inferring functionality-revealing interaction contexts. fSIM-NET is complemented by a generative network (iGEN-NET) and a segmentation network (iSEG-NET). iGEN-NET takes a

single voxelized 3D object with a functionality label and synthesizes a voxelized surround, i.e., the interaction context which visually demonstrates the corresponding functionality. iSEG-NET further separates the interacting objects into different groups according to their interaction types.

CCS Concepts: • **Computing methodologies** → **Shape analysis**;

Additional Key Words and Phrases: 3D object functionality, interaction context, deep convolutional neural network, functionality prediction, synthesis

## ACM Reference Format:

Ruizhen Hu, Zihao Yan, Jingwen Zhang, Oliver van Kaick, Ariel Shamir, Hao Zhang, and Hui Huang. 2018. Predictive and Generative Neural Networks for Object Functionality. *ACM Trans. Graph.* 37, 4, Article 151 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201287>

## 1 INTRODUCTION

In recent years, functionality analysis of 3D shapes has gained attention as a means to understand and manipulate 3D environments. It has been argued that the recognition and categorization of object and scene data are mainly based on their functionality [Greene et al. 2016; Stark and Bowyer 1991]. Even though functionalities of objects can be interpreted in many ways, most of them involve some form of *interactions* between two entities: one that provides the functionality and one that “consumes” it.

Humans can predict the functionality of an object even without any surroundings, since their knowledge and experience would allow them to “hallucinate” the interaction or usage scenarios involving the object. The main question we pose is whether a machine

\*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Ruizhen Hu, College of Computer Science & Software Engineering, Shenzhen University, ruizhen.hu@gmail.com; Zihao Yan, Shenzhen University; Jingwen Zhang, Shenzhen University; Oliver van Kaick, Carleton University; Ariel Shamir, The Interdisciplinary Center; Hao Zhang, Simon Fraser University; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART151 \$15.00

<https://doi.org/10.1145/3197517.3201287>

# P2P-NET: Bidirectional Point Displacement Net for Shape Transform

KANGXUE YIN, Simon Fraser University

HUI HUANG\*, Shenzhen University

DANIEL COHEN-OR, Shenzhen University and Tel Aviv University

HAO ZHANG, Simon Fraser University

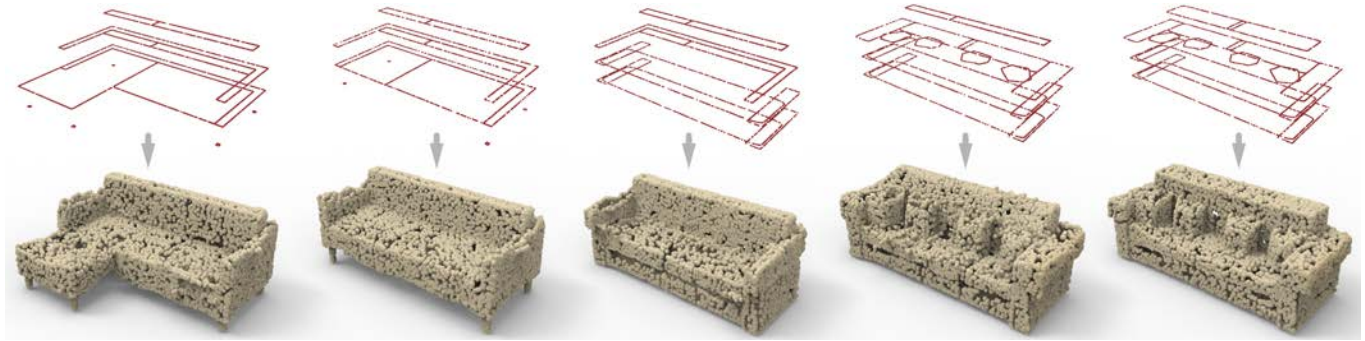


Fig. 1. We develop a general-purpose deep neural network which learns geometric transformations between point sets, e.g., from cross-sectional profiles to 3D shapes, as shown. User can edit the profiles to create an interpolating sequence (top). Our network transforms all of them into point-based 3D shapes.

We introduce P2P-NET, a *general-purpose* deep neural network which learns geometric transformations between point-based shape representations from two domains, e.g., meso-skeletons and surfaces, partial and complete scans, etc. The architecture of the P2P-NET is that of a *bi-directional point displacement* network, which transforms a source point set to a prediction of the target point set with the same cardinality, and vice versa, by applying *point-wise* displacement vectors learned from data. P2P-NET is trained on *paired* shapes from the source and target domains, but *without* relying on point-to-point correspondences between the source and target point sets. The training loss combines two uni-directional geometric losses, each enforcing a *shape-wise* similarity between the predicted and the target point sets, and a cross-regularization term to encourage consistency between displacement vectors going in opposite directions. We develop and present several different applications enabled by our general-purpose bidirectional P2P-NET to highlight the effectiveness, versatility, and potential of our network in solving a variety of point-based shape transformation problems.

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Shape modeling**; **Shape analysis**;

\*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Kangxue Yin, Simon Fraser University; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University; Daniel Cohen-Or, Shenzhen University, Tel Aviv University; Hao Zhang, Simon Fraser University.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART152 \$15.00

<https://doi.org/10.1145/3197517.3201288>

Additional Key Words and Phrases: Point cloud processing, deep neural network, point-wise displacement, point set transform

## ACM Reference Format:

Kangxue Yin, Hui Huang, Daniel Cohen-Or, and Hao Zhang. 2018. P2P-NET: Bidirectional Point Displacement Net for Shape Transform. *ACM Trans. Graph.* 37, 4, Article 152 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201288>

## 1 INTRODUCTION

Point primitives and point-based processing have attracted considerable interest from the computer graphics community for many years [Gross and Pfister 2007]. As a fundamental shape representation, point sets are more compact than voxels and more flexible than polygon meshes. They are immediately available as the default output from most 3D shape acquisition devices. Recently, deep neural networks have been designed to learn global and multi-scale point set features for shape classification and segmentation [Qi et al. 2017a,b], as well as geometry (e.g., normal or curvature) estimation [Guerrero et al. 2017]. Image-driven generative models have also been trained to reconstruct point-based 3D object representations from single or multi-view images [Fan et al. 2017; Lin et al. 2018; Lun et al. 2017].

In this paper, we are interested in exploring how deep neural networks can benefit a new class of problems in point-based graphics: *geometric transformations* between shapes represented by point sets. These shape transforms span a wide spectrum of applications. Some examples include transforms between shape skeletons and surfaces, incomplete and completed object scans, 2D contours and 3D shapes, and simplified and detailed surfaces. Our goal is to develop a *general-purpose* neural network that is capable of learning geometric transformations between point sets from two domains.



# Box Cutter: Atlas Refinement for Efficient Packing via Void Elimination

MAX LIMPER, Fraunhofer IGD / TU Darmstadt, Germany

NICHOLAS VINING, University of British Columbia, Canada

ALLA SHEFFER, University of British Columbia, Canada

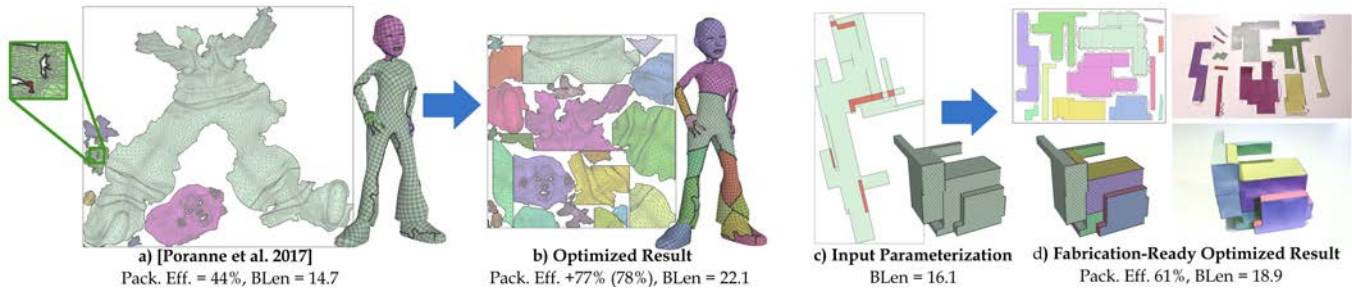


Fig. 1. Traditional texture atlas generation frameworks, such as [Poranne et al. 2017] (a), produce results which can have low packing efficiency and are not necessarily bijective (see inset). We produce an overlap-free atlas with the same parametric distortion and significantly higher packing efficiency by strategically segmenting and repacking the original charts directly in 2D space (b). We use the same framework to efficiently pack 2D patterns for 3D fabrication (c-e).

Packed atlases, consisting of 2D parameterized charts, are ubiquitously used to store surface signals such as texture or normals. Tight packing is similarly used to arrange and cut-out 2D panels for fabrication from sheet materials. *Packing efficiency*, or the ratio between the areas of the packed atlas and its bounding box, significantly impacts downstream applications.

We propose *Box Cutter*, a new method for optimizing packing efficiency suitable for both settings. Our algorithm improves packing efficiency without changing distortion by strategically cutting and repacking the atlas charts or panels. It preserves the local mapping between the 3D surface and the atlas charts and retains global mapping continuity across the newly formed cuts. We balance packing efficiency improvement against increase in chart boundary length and enable users to directly control the acceptable amount of boundary elongation. While the problem we address is NP-hard, we provide an effective practical solution by iteratively detecting large rectangular empty spaces, or *void boxes*, in the current atlas packing and eliminating them by first refining the atlas using strategically placed axis-aligned cuts and then repacking the refined charts. We repeat this process until no further improvement is possible, or until the desired balance between packing improvement and boundary elongation is achieved. Packed chart atlases are only useful for the applications we address if their charts are overlap-free; yet many popular parameterization methods, used as-is, produce atlases with global overlaps. Our pre-processing step eliminates all input overlaps while explicitly minimizing the boundary length of the resulting overlap-free charts. We demonstrate our combined strategy on a large range of input atlases produced by diverse parameterization methods, as well as on multiple

sets of 2D fabrication panels. Our framework dramatically improves the output packing efficiency on all inputs; for instance with boundary length increase capped at 50% we improve packing efficiency by 68% on average.

CCS Concepts: • **Computing methodologies** → *Texturing; Mesh geometry models; Parametric curve and surface models*;

Additional Key Words and Phrases: Texture Atlas, 2D Patterns, Packing, Packing Efficiency

## ACM Reference Format:

Max Limper, Nicholas Vining, and Alla Sheffer. 2018. Box Cutter: Atlas Refinement for Efficient Packing via Void Elimination. *ACM Trans. Graph.* 37, 4, Article 153 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201328>

## 1 INTRODUCTION

2D parameterized chart atlases are ubiquitously used to store surface signals such as colors, textures, or normals [Hormann et al. 2007; Sheffer et al. 2007]. While many popular parameterization methods produce irregularly shaped atlas charts, the actual signals they encode are typically stored as 2D rectangular images defined over the bounding box of the packed atlas to allow for efficient access and GPU processing. Any space within this box that is not occupied by the atlas charts (Figure 1a) is wasted as it contains no actual signal. Wasting significant amounts of both memory and disk space is a concern for asset-heavy real-time applications such as games. Space waste considerations are also of concern for applications that use 2D panels for fabrication: when panels are cut out from flat material sheets, the material within the bounding box of the cut panels is often too fragmented to re-use for other needs and is consequently discarded [Koo et al. 2017] (Figure 1c). Our *Box Cutter* algorithm reduces material and memory waste by refining and repacking input atlases to explicitly maximize their *packing efficiency*, or the ratio between the sum of areas of the atlas charts and the area of their bounding box (Figure 1bd).

We achieve this goal by strategically cutting the existing charts to facilitate more efficient packing. In selecting the cuts, we seek to

Authors' addresses: Max Limper, [max.limper@siggraph.org](mailto:max.limper@siggraph.org), Fraunhofer IGD / TU Darmstadt, Fraunhoferstr. 5, Darmstadt, 64283, Germany; Nicholas Vining, [nvining@cs.ubc.ca](mailto:nvining@cs.ubc.ca), University of British Columbia, 201-2366 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada; Alla Sheffer, [sheffa@cs.ubc.ca](mailto:sheffa@cs.ubc.ca), University of British Columbia, 201-2366 Main Mall, Vancouver, British Columbia, V6T 1Z4, Canada.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2018/8-ART153 \$15.00

<https://doi.org/10.1145/3197517.3201328>

# Gradient-Domain Processing within a Texture Atlas

FABIÁN PRADA and MISHA KAZHDAN, Johns Hopkins University, USA  
 MING CHUANG, PerceptIn Inc., USA  
 HUGUES HOPPE, Google Inc., USA

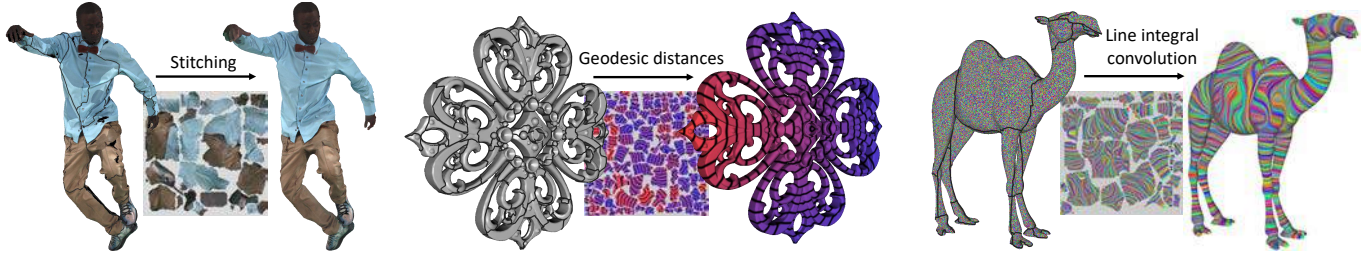


Fig. 1. We process surface signals directly in the texture atlas domain, thereby exploiting the regularity of the 2D grid sampling. Example applications include multiview stitching, computation of geodesic distance maps, and curvature-guided line integral convolution. (Black curves indicate chart boundaries.)

Processing signals on surfaces often involves resampling the signal over the vertices of a dense mesh and applying mesh-based filtering operators. We present a framework to process a signal directly in a texture atlas domain. The benefits are twofold: avoiding resampling degradation and exploiting the regularity of the texture image grid. The main challenges are to preserve continuity across atlas chart boundaries and to adapt differential operators to the non-uniform parameterization. We introduce a novel function space and multigrid solver that jointly enable robust, interactive, and geometry-aware signal processing. We demonstrate our approach using several applications including smoothing and sharpening, multiview stitching, geodesic distance computation, and line integral convolution.

CCS Concepts: • **Computing methodologies** → *Computer graphics*;

Additional Key Words and Phrases: signal processing, Laplacian filtering, mesh parameterization, multigrid solver, domain decomposition

## ACM Reference Format:

Fabián Prada, Misha Kazhdan, Ming Chuang, and Hugues Hoppe. 2018. Gradient-Domain Processing within a Texture Atlas. *ACM Trans. Graph.* 37, 4, Article 154 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201317>

## 1 INTRODUCTION

In computer graphics, detailed surface fields are commonly stored using a *texture atlas parameterization*. The approach is to partition a surface mesh into chart regions, flatten each chart onto a polygon, and pack these polygons into a rectangular atlas domain by

recording texture coordinates at vertices. The resulting texture map efficiently captures high-resolution content and is natively supported by hardware rasterization. A key benefit is that texels lie on a regular image grid, enabling efficient random-access, strong memory coherence, and massive parallelism.

As reviewed in Section 2, most techniques for processing signals on surfaces involve sampling the signal over a dense triangle mesh and defining a discretization of the Laplace operator that adapts to the nonuniform structure and geometry of mesh neighborhoods. We instead explore a framework to perform gradient-domain processing *directly* in a texture atlas domain, thereby (1) eliminating the need for resampling and (2) mapping computation to a regular 2D grid.

**Goal.** Our gradient-domain processing objective is designed to solve a broad class of problems. Given a texture-atlased triangle mesh with metric  $h$ , and given target texture values  $\psi$ , target texture differential  $\omega$ , and a screening weight  $\alpha \geq 0$  balancing fidelity to  $\psi$  and  $\omega$ , the goal is to find the texture values  $\phi$  that minimize the energy

$$E(\phi; h, \alpha, \psi, \omega) = \alpha \cdot \|\phi - \psi\|_h^2 + \|d\phi - \omega\|_h^2. \quad (1)$$

Formulating this least squares minimization over the texel values in a texture atlas poses several challenges:

- Using standard bilinear interpolation, texture maps represent functions that do not (in general) align across chart boundaries. As a result, continuity can only be enforced by constraining the texture signal to have constant value along the seams.
- Evaluating the texture near chart boundaries requires the use of both *interior* and *exterior* texels. Because exterior texels are not associated with positions on the surface, defining discrete derivatives across chart boundaries is non-trivial.
- Although texels lie on a uniform grid, their corresponding locations on the surface are distorted by the parameterization. The nonuniform metric must be taken into account.

Authors' addresses: Fabián Prada; Misha Kazhdan, Johns Hopkins University, 3400 N Charles St, Baltimore, MD, 21218, USA, [fabianprada@gmail.com](mailto:fabianprada@gmail.com), [misha@cs.jhu.edu](mailto:misha@cs.jhu.edu); Ming Chuang, PerceptIn Inc. 4633 Old Ironsides Dr, Snata Clara, CA, 95054, USA, [mingchuang82@gmail.com](mailto:mingchuang82@gmail.com); Hugues Hoppe, Google Inc. 601 N 34th St, Seattle, WA, 98103, USA, [hhoppe@gmail.com](mailto:hhoppe@gmail.com).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM. 0730-0301/2018/8-ART154 \$15.00 <https://doi.org/10.1145/3197517.3201317>

# Generalized Motorcycle Graphs for Imperfect Quad-Dominant Meshes

NICO SCHERTLER, TU Dresden, Germany

DANIELE PANOZZO, New York University, USA

STEFAN GUMHOLD, TU Dresden, Germany

MARCO TARINI, Università degli Studi di Milano, Italy and ISTI - CNR, Italy

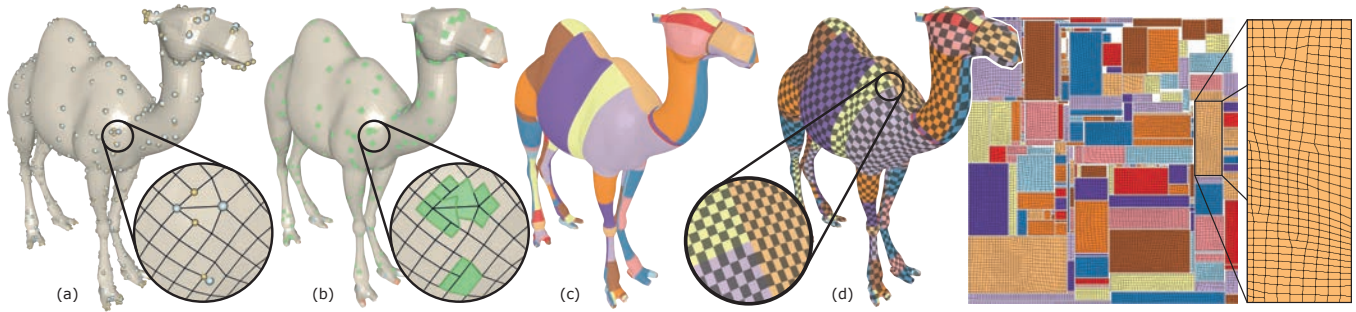


Fig. 1. Our method takes as input a semi-regular quad-dominant mesh (a – singularities marked with spheres) and produces a global parametrization. For that purpose, we evaluate the effect of singularities on the mesh topology with the help of *fenced regions* (b). We use this information to calculate a Generalized Motorcycle Graph (c), whose patches serve as rectangular domains for the parametrization (d). The parametrizations of adjacent patches are aligned to each other on a majority of cuts (see close-up), which allows to make them invisible in texturing applications. The rectangular shape of patches in the 2D parametric domain allows highly efficient packing of the texture. Model courtesy of INRIA (AIM@SHAPE).

We introduce a practical pipeline to create UV T-layouts for real-world quad dominant semi-regular meshes. Our algorithm creates large rectangular patches by relaxing the notion of motorcycle graphs and making it insensitive to local irregularities in the mesh structure such as non-quad elements, redundant irregular vertices, T-junctions, and others. Each surface patch, which can contain multiple singularities and/or polygonal elements, is mapped to an axis-aligned rectangle, leading to a simple and efficient UV layout, which is ideal for texture mapping (allowing for mipmapping and artifact-free bilinear interpolation). We demonstrate that our algorithm is an ideal solution for both recent semi-regular, quad-dominant meshing methods, and for the low-poly meshes typically used in games and movies.

CCS Concepts: • **Computing methodologies** → **Texturing**;

Additional Key Words and Phrases: Texture Mapping, Motorcycle Graph, Parametrization

## ACM Reference Format:

Nico Schertler, Daniele Panozzo, Stefan Gumhold, and Marco Tarini. 2018. Generalized Motorcycle Graphs for Imperfect Quad-Dominant Meshes. *ACM*

Authors' addresses: Nico Schertler, TU Dresden, Germany; Daniele Panozzo, New York University, USA; Stefan Gumhold, TU Dresden, Germany; Marco Tarini, Università degli Studi di Milano, Varese, Italy, ISTI - CNR, Pisa, Italy.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART155 \$15.00

<https://doi.org/10.1145/3197517.3201389>

*Trans. Graph.* 37, 4, Article 155 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201389>

## 1 INTRODUCTION

Quad-dominant semi-structured meshes, i.e. meshes that are predominantly composed of quadrilateral faces and regular vertices, are ubiquitous in computer graphics: they are the de-facto standard in the visual effects and computer animation industry and are also often used in most interactive applications. However, *clean* structured meshes are expensive to create. The number of irregular vertices is often too large and most pipelines do not support common small imperfections, requiring manual cleaning. This slows down existing content creation pipelines and prevents the direct usage of scanned models since reconstruction methods rarely create models adhering to these strict guidelines, even after an automatic cleanup.

We propose a practical way to make imperfect polygonal meshes more directly usable in downstream applications, by allowing the computation of a valid UV-layout robust to redundant irregular vertices as well as to quad mesh imperfections such as small holes, non-quadrilateral faces, T-junctions, small handles and tunnels, and non two-manifold configurations.

Our main contribution is the definition of a generalization of *Motorcycle Graphs* (MCG) [Eppstein et al. 2008]: our construction is identical on clean, highly regular quad meshes but gracefully handles imperfections and redundant irregular vertices (Section 1.1), always producing a valid segmentation into rectangular UV patches. The key idea is to identify a set of regions on the mesh which are to be considered equivalent to completely regular and clean grids for the purpose of the MCG algorithm, in spite of imperfections.



# Variational Surface Cutting

NICHOLAS SHARP and KEENAN CRANE, Carnegie Mellon University



Fig. 1. Our flow evolves a given cut toward one that yields lower area distortion when the surface is flattened. Here we start with a cut through the equator and flow to cuts of three different lengths (visualized as seams on the tennis ball), resulting in progressively easier flattenings into the plane.

This paper develops a global variational approach to cutting curved surfaces so that they can be flattened into the plane with low metric distortion. Such cuts are a critical component in a variety of algorithms that seek to parameterize surfaces over flat domains, or fabricate structures from flat materials. Rather than evaluate the quality of a cut solely based on properties of the curve itself (e.g., its length or curvature), we formulate a flow that directly optimizes the distortion induced by cutting and flattening. Notably, we do not have to explicitly parameterize the surface in order to evaluate the cost of a cut, but can instead integrate a simple evolution equation defined on the cut curve itself. We arrive at this flow via a novel application of *shape derivatives* to the *Yamabe equation* from conformal geometry. We then develop an Eulerian numerical integrator on triangulated surfaces, which does not restrict cuts to mesh edges and can incorporate user-defined data such as importance or occlusion. The resulting cut curves can be used to drive distortion to arbitrarily low levels, and have a very different character from cuts obtained via purely discrete formulations. We briefly explore potential applications to computational design, as well as connections to space filling curves and the problem of uniform heat distribution.

CCS Concepts: • **Computing methodologies** → **Mesh geometry models**; • **Mathematics of computing** → *Continuous optimization*;

Additional Key Words and Phrases: discrete differential geometry, geometry processing

Authors' address: Nicholas Sharp; Keenan Crane, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA, 15213.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

0730-0301/2018/8-ART156 \$15.00  
<https://doi.org/10.1145/3197517.3201356>

## ACM Reference Format:

Nicholas Sharp and Keenan Crane. 2018. Variational Surface Cutting. *ACM Trans. Graph.* 37, 4, Article 156 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201356>

## 1 INTRODUCTION

Which cut allows a curved surface to be most easily flattened? In general, surfaces cannot be flattened without some distortion of areas or angles—a classic example is maps of the Earth, which invariably distort the relative shape or size of land masses. Broader interest in cutting and flattening curved surfaces arises in a variety of problems in digital geometry processing, scientific computing, and computational design, where distortion affects things like signal quality, numerical accuracy, and material stress during manufacturing. The general strategy for reducing distortion is to make cuts in the surface, much as a designer might accommodate curvature by incorporating seams into a pattern for a garment. However, the question of *where* to place cuts remains an ongoing challenge.

A common idea is to view cutting as an inherently discrete problem: find the best cut along edges of a triangulation. However, these edges may have no meaning in a context like physical manufacturing, where the triangulation is merely a proxy for a smooth surface. In this paper we instead start in the smooth setting, and show that the distortion of a cut can be minimized by integrating a remarkably simple flow. In particular, if  $n$  is the unit normal of a curve  $\gamma$ , then we simply need to evolve  $\gamma$  according to the equation

$$\frac{d}{dt}\gamma = \frac{\partial u}{\partial n}^2 n, \quad (1)$$

where  $u$  is the logarithm of the scale distortion induced by cutting the surface along  $\gamma$  and applying a conformal (i.e., angle-preserving)

# 3D Printing Spatially Varying Color and Translucency

ALAN BRUNTON, CAN ATES ARIKAN, and TEJAS MADAN TANKSALE, Fraunhofer IGD  
PHILIPP URBAN, Fraunhofer IGD and Norwegian University of Science and Technology NTNU



Fig. 1. Design control and perceptual reproduction of material translucency for full color 3D printing: (a) a printed anatomy model with opaque and translucent parts, revealing internal structures; (b) a polished transparent Lion with a red mane and 3D logo inside; (c) a transparent lizard with colored scales illuminated from below by a checkerboard pattern; (d) a close-up (see Figure 12 for the full view) of wax and a 3D print reproducing the wax's translucent appearance.

We present an efficient and scalable pipeline for fabricating full-colored objects with spatially-varying translucency from practical and accessible input data via multi-material 3D printing. Observing that the costs associated with BSSRDF measurement and processing are high, the range of 3D printable BSSRDFs are severely limited, and that the human visual system relies only on simple high-level cues to perceive translucency, we propose a method based on reproducing perceptual translucency cues. The input to our pipeline is an RGBA signal defined on the surface of an object, making our approach accessible and practical for designers. We propose a framework for extending standard color management and profiling to combined color and translucency management using a gamut correspondence strategy we call *opaque relative processing*. We present an efficient streaming method to compute voxel-level material arrangements, achieving both realistic reproduction of measured translucent materials and artistic effects involving multiple fully or partially transparent geometries.

CCS Concepts: • **Computing methodologies** → **Perception**; *Graphics file formats*;

Additional Key Words and Phrases: 3D printing, metamerism, translucency, color, distance field

## ACM Reference Format:

Alan Brunton, Can Ates Arikan, Tejas Madan Tanksale, and Philipp Urban. 2018. 3D Printing Spatially Varying Color and Translucency. *ACM Trans. Graph.* 37, 4, Article 157 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201349>

Authors' addresses: Alan Brunton, [alan.brunton@igd.fraunhofer.de](mailto:alan.brunton@igd.fraunhofer.de); Can Ates Arikan, [arikan.canates@gmail.com](mailto:arikan.canates@gmail.com); Tejas Madan Tanksale, [tejas.madan.tanksale@igd.fraunhofer.de](mailto:tejas.madan.tanksale@igd.fraunhofer.de); Philipp Urban, [philipp.urban@igd.fraunhofer.de](mailto:philipp.urban@igd.fraunhofer.de) Fraunhofer IGD Norwegian University of Science and Technology NTNU.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART157 \$15.00

<https://doi.org/10.1145/3197517.3201349>

## 1 INTRODUCTION

We present an efficient and scalable pipeline for fabricating full-colored objects with spatially-varying translucency from practical and accessible input data via multi-material 3D printing. Color and translucency are encoded by an RGBA signal on the object's surface (i.e. an RGBA texture or per-vertex RGBA attributes), which is supported by various 3D file formats, making our method applicable on many existing models and facilitating model design and creation via existing 3D modeling software.

In contrast to previous approaches focusing only on albedo color fabrication [Babaei et al. 2017; Brunton et al. 2015; Elek et al. 2017; Panozzo et al. 2015; Schüller et al. 2016] or only on translucency fabrication [Dong et al. 2010; Hašan et al. 2010], our method minimizes translucency errors subject to accurate color reproduction. It allows fabricating full-color objects possessing almost opaque, translucent, and fully transparent parts, with both smooth and abrupt changes between these regions. Since many common materials are partially permeable to light such as wax, skin, teeth, cheese, fish, or stone, our method substantially enhances the degree of realism of 3D prints.

Light transport between any two rays that hit the object's surface is phenomenologically modeled by the spatially-varying Bidirectional Surface Scattering Reflectance Distribution Function (BSSRDF) [Niedermeyer et al. 1977], intrinsically accounting for spectrally varying absorption and scattering, and phase modulation. A full physical description is thus very high dimension, and multiple works have sought low-dimensional physical approximations [Jensen et al. 2001; Song et al. 2009]. However, by working in the perceptual domain we can significantly reduce the dimensionality even further.

In addition, measuring BSSRDFs [Goessele et al. 2004; Peers et al. 2006; Tong et al. 2005] is expensive in terms of system complexity, acquisition time, storage and processing of captured data. This is true for both natural materials we would like to reproduce, and the output of a fabrication system to characterize the set of BSSRDFs it can produce. There exists no standard format for storing a BSSRDF.

Multi-material 3D printing systems are inherently limited in terms of the set BSSRDFs they can reproduce. This is due to the limited

# Fabricating Reflectors for Displaying Multiple Images

KAISEI SAKURAI, DWANGO Co., Ltd., Dwango CG Research

YOSHINORI DOBASHI, Hokkaido University and Dwango CG Research

KEI IWASAKI, Wakayama University and Dwango CG Research

TOMOYUKI NISHITA, Dwango CG Research and Hiroshima Shudo University

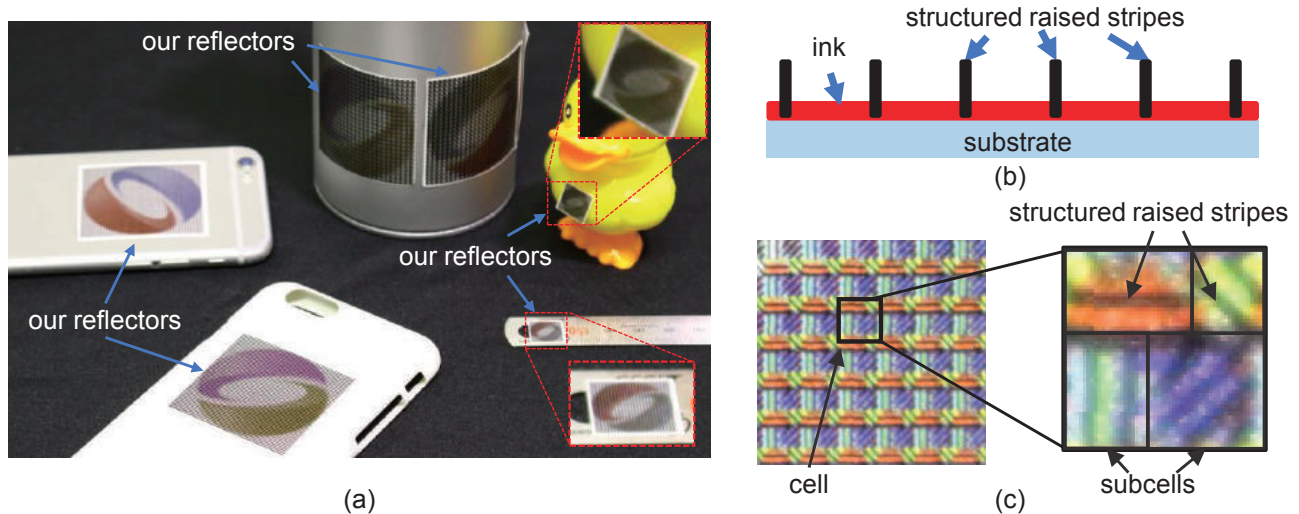


Fig. 1. (a) Examples of our reflectors printed on everyday personal objects: a pen holder, smartphone cases, toy ducks, and a ruler. The reflectors on the smartphone cases are directly printed by a standard UV printer. Other reflectors are printed on stickers and pasted onto the objects. All the reflectors are designed to display SIGGRAPH logo with different colors depending on the viewing direction (see Fig. 6 for input images). (b) Microstructure of the printed pattern consisting of raised stripes on an ink layer. (c) Close-up photograph of printed patterns.

A great deal of attention has been devoted to the fabrication of reflectors that can display different color images when viewed from different directions not only in industry but also for the arts. Although such reflectors have previously been successfully fabricated, the number of images displayed has been limited to two or they suffer from ghosting artifacts where mixed images appear. Furthermore, the previous methods need special hardware and/or materials to fabricate the reflectors. Thus, those techniques are not suitable for printing reflectors on everyday personal objects made of different materials, such as name cards, letter sheets, envelopes, and plastic cases. To overcome these limitations, we propose a method for fabricating reflectors using a standard ultraviolet printer (UV printer). UV printer can render a specified 2D color pattern on an arbitrary material and by overprinting the printed pattern can be raised, that is, the printed pattern

Authors' addresses: Kaisei Sakurai, DWANGO Co., Ltd., Dwango CG Research, kaisei\_sakurai@dwango.co.jp; Yoshinori Dobashi, Hokkaido University, Dwango CG Research, doba@ime.ist.hokudai.ac.jp; Kei Iwasaki, Wakayama University, Dwango CG Research, iwasaki@sys.wakayama-u.ac.jp; Tomoyuki Nishita, Dwango CG Research, Hiroshima Shudo University, nishita@shudo-u.ac.jp.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART158 \$15.00

<https://doi.org/10.1145/3197517.3201400>

becomes a microstructure having color and height. We propose using these microstructures to formulate a method for designing spatially varying reflections that can display different target images when viewed from different directions. The microstructure is calculated by minimizing an objective function that measures the differences between the intensities of the light reflected from the reflector and that of the target image. We show several fabricated reflectors to demonstrate the usefulness of the proposed method.

CCS Concepts: • **Computing methodologies** → *Image manipulation*; *Reflectance modeling*;

Additional Key Words and Phrases: digital fabrication, reflector, ultraviolet printer, spatially varying anisotropic BRDF

## ACM Reference Format:

Kaisei Sakurai, Yoshinori Dobashi, Kei Iwasaki, and Tomoyuki Nishita. 2018. Fabricating Reflectors for Displaying Multiple Images. *ACM Trans. Graph.* 37, 4, Article 158 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201400>

## 1 INTRODUCTION

To address the demand for increasing variation in industrial design, a number of techniques for displaying images have been developed, such as reflectors [Regg et al. 2010; Snelgrove et al. 2013; Weyrich et al. 2007], relief sculptures [Alexa and Matusik 2010; Belhumeur et al. 1999; Mitra and Pauly 2009; Schüller et al. 2014; Song et al. 2007; Sýkora et al. 2014], and lenses [Baran et al. 2012; Papas et al. 2012;



# Computational Design of Nanostructural Color for Additive Manufacturing

THOMAS AUZINGER, Institute of Science and Technology Austria, Austria

WOLFGANG HEIDRICH, King Abdullah University of Science and Technology, Kingdom of Saudi Arabia

BERND BICKEL, Institute of Science and Technology Austria, Austria



Fig. 1. (Center) Microscope image of a structurally colored surface with the corresponding nanostructures of the fore- and background depicted left and right of it. (Top) The cross-sectional profiles of both transparent structures were designed with our method using full electromagnetic simulation. (Bottom) Micrographs of the fabricated nanostructures with a length scale of 5  $\mu\text{m}$  across the image diagonal.

Additive manufacturing has recently seen drastic improvements in resolution, making it now possible to fabricate features at scales of hundreds or even dozens of nanometers, which previously required very expensive lithographic methods. As a result, additive manufacturing now seems poised for optical applications, including those relevant to computer graphics, such as material design, as well as display and imaging applications.

In this work, we explore the use of additive manufacturing for generating structural colors, where the structures are designed using a fabrication-aware optimization process. This requires a combination of full-wave simulation, a feasible parameterization of the design space, and a tailored optimization procedure. Many of these components should be re-usable for the design of other optical structures at this scale.

We show initial results of material samples fabricated based on our designs. While these suffer from the prototype character of state-of-the-art fabrication hardware, we believe they clearly demonstrate the potential of additive nanofabrication for structural colors and other graphics applications.

Authors' addresses: Thomas Auzinger, Institute of Science and Technology Austria, Am Campus 1, 3400, Klosterneuburg, Austria, thomas.auzinger@ist.ac.at; Wolfgang Heidrich, King Abdullah University of Science and Technology, Al Khwarizmi Building 1, 23955-6900, Thuwal, Kingdom of Saudi Arabia, wolfgang.heidrich@kaust.edu.sa; Bernd Bickel, Institute of Science and Technology Austria, Am Campus 1, 3400, Klosterneuburg, Austria, bernd.bickel@ist.ac.at.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

0730-0301/2018/8-ART159 \$15.00

<https://doi.org/10.1145/3197517.3201376>

CCS Concepts: • **Applied computing** → **Computer-aided manufacturing**; • **Hardware** → *Emerging optical and photonic technologies*;

Additional Key Words and Phrases: structural colorization, appearance, multiphoton lithography, direct laser writing, computational fabrication, computational design, shape optimization, FDTD, diffraction, Nanoscribe

## ACM Reference Format:

Thomas Auzinger, Wolfgang Heidrich, and Bernd Bickel. 2018. Computational Design of Nanostructural Color for Additive Manufacturing. *ACM Trans. Graph.* 37, 4, Article 159 (August 2018), 16 pages. <https://doi.org/10.1145/3197517.3201376>

## 1 INTRODUCTION

In contrast to pigment-based colors – where spectral components of incident light are absorbed by the pigment material – *structural colorization* arises from the interaction of light with micro- and nanostructures. Variations of the refractive index at the length scale of the light's wavelength cause interference effects that lead to fascinating optical effects such as iridescence. Designing such colorization is challenging, however, since the wave nature of light has to be taken into account when simulating its interplay with nanostructures. Furthermore, realizing such structures requires elaborate methods from nanofabrication. Owing to these complications, linking computational methods and advanced manufacturing is essential for exploring structural colorization, but so far has received little attention from the scientific community, compared to recent progress in designing structures for elastic, auxetic, translucent, or reflective materials [Bermano et al. 2017; Hullin et al. 2013].

In recent years, an advanced additive manufacturing technology has reached commercial maturity, which permits the fabrication of sub-micrometer structures. *Multiphoton lithography* – also known

# Skaterbots: Optimization-based design and motion synthesis for robotic creatures with legs and wheels

MORITZ GEILINGER, ETH Zurich

ROI PORANNE, ETH Zurich

RUTA DESAI, Carnegie Mellon University

BERNHARD THOMASZEWSKI, Université de Montréal

STELIAN COROS, ETH Zurich



Fig. 1. Robotic creatures created with our computational design system employ arbitrary arrangements of legs and wheels to locomote.

We present a computation-driven approach to design optimization and motion synthesis for robotic creatures that locomote using arbitrary arrangements of legs and wheels. Through an intuitive interface, designers first create unique robots by combining different types of servomotors, 3D printable connectors, wheels and feet in a mix-and-match manner. With the resulting robot as input, a novel trajectory optimization formulation generates walking, rolling, gliding and skating motions. These motions emerge naturally based on the components used to design each individual robot. We exploit the particular structure of our formulation and make targeted simplifications to significantly accelerate the underlying numerical solver without compromising quality. This allows designers to interactively choreograph stable, physically-valid motions that are agile and compelling. We furthermore develop a suite of user-guided, semi-automatic, and fully-automatic optimization tools that enable motion-aware edits of the robot's physical structure. We demonstrate the efficacy of our design methodology by creating a diverse array of hybrid legged/wheeled mobile robots which we validate using physics simulation and through fabricated prototypes.

CCS Concepts: • **Computer graphics** → **Computational geometry and object modeling**; *Physically based modeling*;

Additional Key Words and Phrases: computational design, mobile robots, animation, physical characters

Authors' addresses: Moritz Geilinger, ETH Zurich; Roi Poranne, ETH Zurich; Ruta Desai, Carnegie Mellon University; Bernhard Thomaszewski, Université de Montréal; Stelian Coros, ETH Zurich.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2018 Copyright held by the owner/author(s).

0730-0301/2018/8-ART160

<https://doi.org/10.1145/3197517.3201368>

## ACM Reference Format:

Moritz Geilinger, Roi Poranne, Ruta Desai, Bernhard Thomaszewski, and Stelian Coros. 2018. Skaterbots: Optimization-based design and motion synthesis for robotic creatures with legs and wheels. *ACM Trans. Graph.* 37, 4, Article 160 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201368>

## 1 INTRODUCTION

Whether it is to help with chores, keep us company, or entertain us, personal robots promise to play a central role in our increasingly technology-driven society. Echoing the trend of mass customization and leveraging recent advances in digital fabrication, our long term goal is to develop algorithmic foundations that will enable these robots to be created on-demand according to the individual needs and preferences of those they serve. In this quest, we join recent research efforts that bridge the fields of animation, fabrication-oriented design and robotics [Bern et al. 2017; Du et al. 2016; Schulz et al. 2017]. Complementing this body of work, we introduce a novel design system for a rich class of mobile robots that employ arbitrary arrangements of legs and wheels for locomotion. Such hybrid robots enjoy the combined versatility of legged and wheeled systems, but they also inherit their compounded challenges: they have many actuated degrees of freedom that need to be precisely coordinated in order to generate motions that are balanced, elegant, and efficient; their kinematics and dynamics are governed by highly non-linear equations; and their motor capabilities and physical design characteristics are inseparably intertwined. For these reasons, creating hybrid mobile robots remains a very difficult and error-prone task.

We present a computation-driven approach to designing, optimizing and synthesizing motions for different breeds of legged/wheeled robots. At the core of our work lies an efficient trajectory optimization formulation tailored to the specific challenges of this class of

# VisemeNet: Audio-Driven Animator-Centric Speech Animation

YANG ZHOU, University of Massachusetts Amherst

ZHAN XU, University of Massachusetts Amherst

CHRIS LANDRETH, University of Toronto

EVANGELOS KALOGERAKIS, University of Massachusetts Amherst

SUBHRANSU MAJI, University of Massachusetts Amherst

KARAN SINGH, University of Toronto

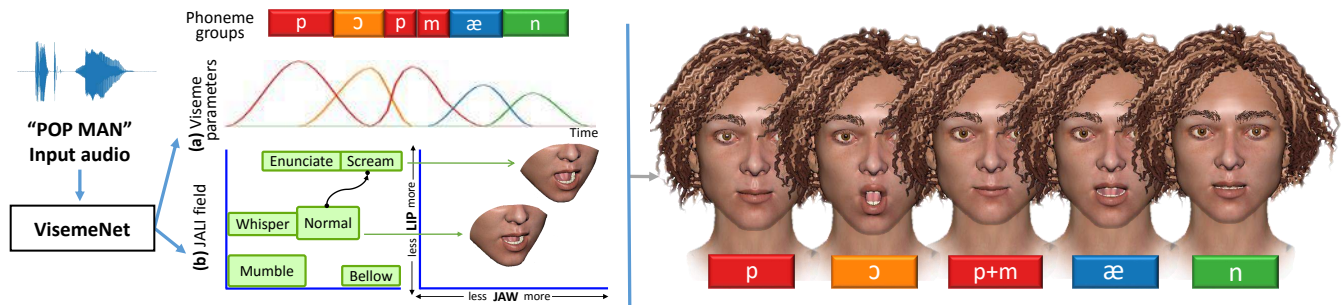


Fig. 1. VisemeNet is a deep-learning approach that uses a 3-stage LSTM network, to predict compact animator-centric viseme curves with proper co-articulation, and speech style parameters, directly from speech audio in near real-time (120ms lag).

We present a novel deep-learning based approach to producing animator-centric speech motion curves that drive a JALI or standard FACS-based production face-rig, directly from input audio. Our three-stage Long Short-Term Memory (LSTM) network architecture is motivated by psycho-linguistic insights: segmenting speech audio into a stream of phonetic-groups is sufficient for viseme construction; speech styles like mumbling or shouting are strongly co-related to the motion of facial landmarks; and animator style is encoded in viseme motion curve profiles. Our contribution is an automatic real-time lip-synchronization from audio solution that integrates seamlessly into existing animation pipelines. We evaluate our results by: cross-validation to ground-truth data; animator critique and edits; visual comparison to recent deep-learning lip-synchronization solutions; and showing our approach to be resilient to diversity in speaker and language.

CCS Concepts: • **Computing methodologies** → **Animation**; *Machine learning algorithms*;

Additional Key Words and Phrases: facial animation, neural networks

## ACM Reference Format:

Yang Zhou, Zhan Xu, Chris Landreth, Evangelos Kalogerakis, Subhransu Maji, and Karan Singh. 2018. VisemeNet: Audio-Driven Animator-Centric Speech Animation. *ACM Trans. Graph.* 37, 4, Article 161 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201292>

Authors' addresses: Yang Zhou, University of Massachusetts Amherst, yangzhou@cs.umass.edu; Zhan Xu, University of Massachusetts Amherst, zhanxu@cs.umass.edu; Chris Landreth, University of Toronto, c.landreth@rogers.com; Evangelos Kalogerakis, University of Massachusetts Amherst, kalo@cs.umass.edu; Subhransu Maji, University of Massachusetts Amherst, smaji@cs.umass.edu; Karan Singh, University of Toronto, karan@dgip.toronto.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART161 \$15.00

<https://doi.org/10.1145/3197517.3201292>

## 1 INTRODUCTION

The importance of realistic computer generated human faces cannot be understated. A diverse set of applications ranging from entertainment (movies and games), medicine (facial therapy and prosthetics) and education (language/speech training and cyber-assistants) are all empowered by the ability to realistically model, simulate and animate human faces. Imperfect emulation of even subtle facial nuance can plunge an animated character into the Uncanny Valley, where the audience loses trust and empathy with the character. Paradoxically, the greater the rendered realism of the character, the less tolerant we are of flaws in its animation [MacDorman et al. 2009]. The expressive animation of speech, unsurprisingly, is a critical component of facial animation that has been the subject of research for almost half-a-century [Bailey et al. 2012].

Typically, keyframing or performance capture are used for high-end commercial animation. Keyframing by professional animators is both expressive and editable, but laborious and prohibitively expensive in time and effort. Performance capture solutions are the opposite; recording a vocal performance is relatively easy, but hard to further edit or refine, and voice actors often have visually inexpressive faces.

Recent approaches however, have shown that there is enough detail in the audio signal itself to produce realistic speech animation. JALI [Edwards et al. 2016] presented a FACS [Ekman and Friesen 1978] and psycho-linguistically inspired face-rig capable of animating a range of speech styles, and an animator-centric procedural lip-synchronization solution using an audio performance and speech transcript as input. Following JALI, a number of concurrent approaches have successfully shown the use of deep learning to capture a mapping between input audio and an animated face [Karras et al. 2017; Suwajanakorn et al. 2017; Taylor et al. 2017]. These approaches however, produce results that, unlike JALI, do not integrate well into a typical animator workflow. This paper addresses the



# High-Fidelity Facial Reflectance and Geometry Inference From an Unconstrained Image

SHUGO YAMAGUCHI\*, Waseda University and USC Institute for Creative Technologies

SHUNSUKE SAITO\*, Pinscreen, University of Southern California, and USC Institute for Creative Technologies

KOKI NAGANO, Pinscreen

YAJIE ZHAO, USC Institute for Creative Technologies

WEIKAI CHEN, USC Institute for Creative Technologies

KYLE OLSZEWSKI, Pinscreen, University of Southern California, and USC Institute for Creative Technologies

SHIGEO MORISHIMA, Waseda University

HAO LI, Pinscreen, University of Southern California, and USC Institute for Creative Technologies

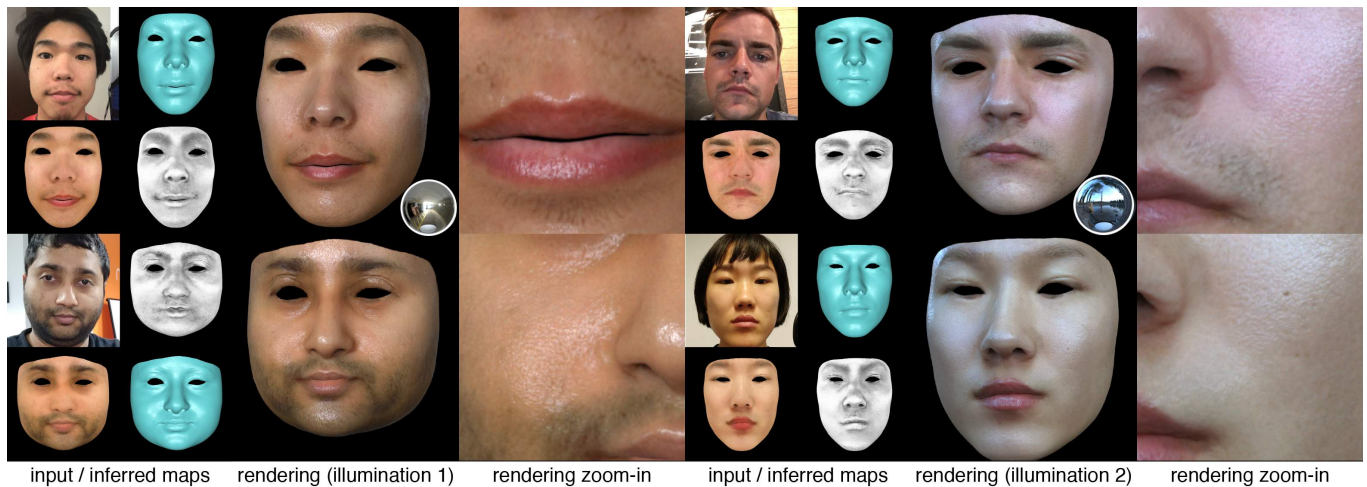


Fig. 1. Our system infers high-fidelity facial reflectance and geometry maps from a single image (diffuse albedo, specular albedo, as well as medium- and high-frequency displacements). These maps can be used for high-fidelity rendering under novel illumination conditions.

We present a deep learning-based technique to infer high-quality facial reflectance and geometry given a single unconstrained image of the subject, which may contain partial occlusions and arbitrary illumination conditions. The reconstructed high-resolution textures, which are generated in

\* indicates equal contribution

Authors' addresses: Shugo Yamaguchi, Waseda University, USC Institute for Creative Technologies, [wasedayshugo@suou.waseda.jp](mailto:wasedayshugo@suou.waseda.jp); Shunsuke Saito, Pinscreen, University of Southern California, USC Institute for Creative Technologies, [shunsuke.saito16@gmail.com](mailto:shunsuke.saito16@gmail.com); Koki Nagano, Pinscreen, [knagano@usc.edu](mailto:knagano@usc.edu); Yajie Zhao, USC Institute for Creative Technologies, [yajie730@gmail.com](mailto:yajie730@gmail.com); Weikai Chen, USC Institute for Creative Technologies, [chenwk891@gmail.com](mailto:chenwk891@gmail.com); Kyle Olszewski, Pinscreen, University of Southern California, USC Institute for Creative Technologies, [olszewsk@usc.edu](mailto:olszewsk@usc.edu); Shigeo Morishima, Waseda University, [shigeo@waseda.jp](mailto:shigeo@waseda.jp); Hao Li, Pinscreen, University of Southern California, USC Institute for Creative Technologies, [hao@hao-li.com](mailto:hao@hao-li.com).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART162 \$15.00

<https://doi.org/10.1145/3197517.3201364>

only a few seconds, include high-resolution skin surface reflectance maps, representing both the diffuse and specular albedo, and medium- and high-frequency displacement maps, thereby allowing us to render compelling digital avatars under novel lighting conditions. To extract this data, we train our deep neural networks with a high-quality skin reflectance and geometry database created with a state-of-the-art multi-view photometric stereo system using polarized gradient illumination. Given the raw facial texture map extracted from the input image, our neural networks synthesize complete reflectance and displacement maps, as well as complete missing regions caused by occlusions. The completed textures exhibit consistent quality throughout the face due to our network architecture, which propagates texture features from the visible region, resulting in high-fidelity details that are consistent with those seen in visible regions. We describe how this highly underconstrained problem is made tractable by dividing the full inference into smaller tasks, which are addressed by dedicated neural networks. We demonstrate the effectiveness of our network design with robust texture completion from images of faces that are largely occluded. With the inferred reflectance and geometry data, we demonstrate the rendering of high-fidelity 3D avatars from a variety of subjects captured under different lighting conditions. In addition, we perform evaluations demonstrating that our method can infer plausible facial reflectance and geometric details comparable to

# Deep Video Portraits

HYEONGWOO KIM, Max Planck Institute for Informatics, Germany

PABLO GARRIDO, Technicolor, France

AYUSH TEWARI and WEIPENG XU, Max Planck Institute for Informatics, Germany

JUSTUS THIES and MATTHIAS NIESSNER, Technical University of Munich, Germany

PATRICK PÉREZ, Technicolor, France

CHRISTIAN RICHARDT, University of Bath, United Kingdom

MICHAEL ZOLLHÖFER, Stanford University, United States of America

CHRISTIAN THEOBALT, Max Planck Institute for Informatics, Germany

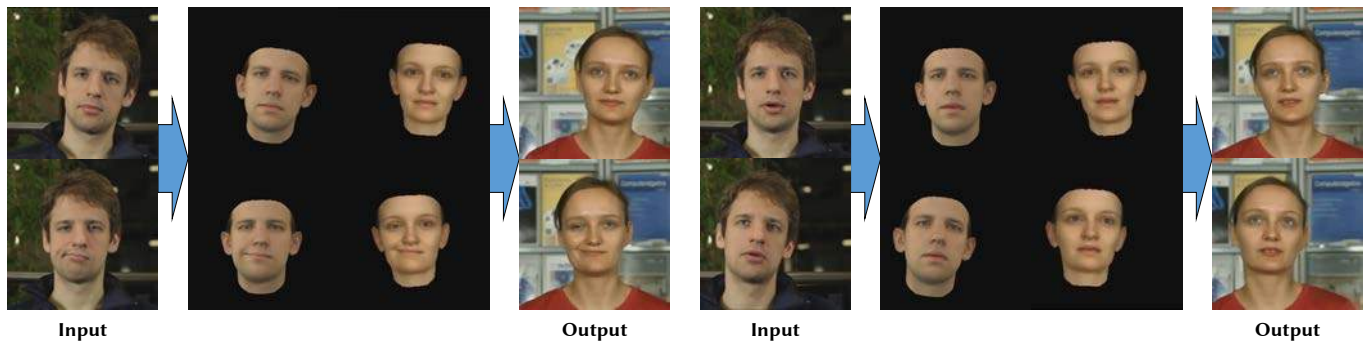


Fig. 1. Unlike current face reenactment approaches that only modify the expression of a target actor in a video, our novel deep video portrait approach enables full control over the target by transferring the rigid head pose, facial expression and eye motion with a high level of photorealism.

We present a novel approach that enables photo-realistic re-animation of portrait videos using only an input video. In contrast to existing approaches that are restricted to manipulations of facial expressions only, we are the first to transfer the full 3D head position, head rotation, face expression, eye gaze, and eye blinking from a source actor to a portrait video of a target actor. The core of our approach is a generative neural network with a novel space-time architecture. The network takes as input synthetic renderings of a parametric face model, based on which it predicts photo-realistic video frames for a given target actor. The realism in this rendering-to-video transfer is achieved by careful adversarial training, and as a result, we can create modified target

videos that mimic the behavior of the synthetically-created input. In order to enable source-to-target video re-animation, we render a synthetic target video with the reconstructed head animation parameters from a source video, and feed it into the trained network – thus taking full control of the target. With the ability to freely recombine source and target parameters, we are able to demonstrate a large variety of video rewrite applications without explicitly modeling hair, body or background. For instance, we can reenact the full head using interactive user-controlled editing, and realize high-fidelity visual dubbing. To demonstrate the high quality of our output, we conduct an extensive series of experiments and evaluations, where for instance a user study shows that our video edits are hard to detect.

Authors' addresses: Hyeonwoo Kim, Max Planck Institute for Informatics, Campus E1.4, Saarbrücken, 66123, Germany, [hyeonwoo.kim@mpi-inf.mpg.de](mailto:hyeonwoo.kim@mpi-inf.mpg.de); Pablo Garrido, Technicolor, 975 Avenue des Champs Blancs, Cesson-Sévigné, 35576, France, [pablo.garrido.adrian@gmail.com](mailto:pablo.garrido.adrian@gmail.com); Ayush Tewari, [atewari@mpi-inf.mpg.de](mailto:atewari@mpi-inf.mpg.de); Weipeng Xu, [w Xu@mpi-inf.mpg.de](mailto:w Xu@mpi-inf.mpg.de), Max Planck Institute for Informatics, Campus E1.4, Saarbrücken, 66123, Germany; Justus Thies, [justus.thies@tum.de](mailto:justus.thies@tum.de); Matthias Nießner, [niessner@tum.de](mailto:niessner@tum.de), Technical University of Munich, Boltzmannstraße 3, Garching, 85748, Germany; Patrick Pérez, Technicolor, 975 Avenue des Champs Blancs, Cesson-Sévigné, 35576, France, [Patrick.Perez@technicolor.com](mailto:Patrick.Perez@technicolor.com); Christian Richardt, University of Bath, Claverton Down, Bath, BA2 7AY, United Kingdom, [christian@richardt.name](mailto:christian@richardt.name); Michael Zollhöfer, Stanford University, 353 Serra Mall, Stanford, CA, 94305, United States of America, [zollhoefer@cs.stanford.edu](mailto:zollhoefer@cs.stanford.edu); Christian Theobalt, Max Planck Institute for Informatics, Campus E1.4, Saarbrücken, 66123, Germany, [theobalt@mpi-inf.mpg.de](mailto:theobalt@mpi-inf.mpg.de).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART163 \$15.00

<https://doi.org/10.1145/3197517.3201283>

CCS Concepts: • **Computing methodologies** → **Computer graphics**; **Neural networks**; *Appearance and texture representations*; *Animation*; *Rendering*;

Additional Key Words and Phrases: Facial Reenactment, Video Portraits, Dubbing, Deep Learning, Conditional GAN, Rendering-to-Video Translation

## ACM Reference Format:

Hyeonwoo Kim, Pablo Garrido, Ayush Tewari, Weipeng Xu, Justus Thies, Matthias Nießner, Patrick Pérez, Christian Richardt, Michael Zollhöfer, and Christian Theobalt. 2018. Deep Video Portraits. *ACM Trans. Graph.* 37, 4, Article 163 (August 2018), 14 pages. <https://doi.org/10.1145/3197517.3201283>

## 1 INTRODUCTION

Synthesizing and editing video portraits, i.e., videos framed to show a person's head and upper body, is an important problem in computer graphics, with applications in video editing and movie post-production, visual effects, visual dubbing, virtual reality, and telepresence, among others. In this paper, we address the problem of

# HeadOn: Real-time Reenactment of Human Portrait Videos

JUSTUS THIES, Technical University of Munich

MICHAEL ZOLLHÖFER, Stanford University

CHRISTIAN THEOBALT, Max-Planck-Institute for Informatics

MARC STAMMINGER, University of Erlangen-Nuremberg

MATTHIAS NIESSNER, Technical University of Munich



Fig. 1. Our novel *HeadOn* approach enables real-time reenactment of upper body motion, head pose, face expression, and eye gaze in human portrait videos. For synthesis of new photo-realistic video content, we employ a novel video-based rendering approach that builds on top of a fully controllable 3D actor model. The person-specific model is constructed from a short RGB-D calibration sequence and is driven by a real-time torso and face tracker.

We propose *HeadOn*, the first real-time source-to-target reenactment approach for complete human portrait videos that enables transfer of torso and head motion, face expression, and eye gaze. Given a short RGB-D video of the target actor, we automatically construct a personalized geometry proxy that embeds a parametric head, eye, and kinematic torso model. A novel real-time reenactment algorithm employs this proxy to photo-realistically map the captured motion from the source actor to the target actor. On top of the coarse geometric proxy, we propose a video-based rendering technique that composites the modified target portrait video via view- and pose-dependent texturing, and creates photo-realistic imagery of the target actor under novel torso and head poses, facial expressions, and gaze directions. To this end, we propose a robust tracking of the face and torso of the source actor. We extensively evaluate our approach and show significant improvements in enabling much greater flexibility in creating realistic reenacted output videos.

CCS Concepts: • **Computing methodologies** → **Computer vision**; **Computer graphics**;

Additional Key Words and Phrases: Reenactment, Face tracking, Video-based Rendering, Real-time

Authors' addresses: Justus Thies, Technical University of Munich, justus.thies@tum.de; Michael Zollhöfer, Stanford University, zollhoefer@cs.stanford.edu; Christian Theobalt, Max-Planck-Institute for Informatics, theobalt@mpi-inf.mpg.de; Marc Stamminger, University of Erlangen-Nuremberg, marc.stamminger@fau.de; Matthias Nießner, Technical University of Munich, niessner@tum.de.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions.acm.org](https://permissions.acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART164

<https://doi.org/10.1145/3197517.3201350>

## ACM Reference Format:

Justus Thies, Michael Zollhöfer, Christian Theobalt, Marc Stamminger, and Matthias Nießner. 2018. *HeadOn: Real-time Reenactment of Human Portrait Videos*. *ACM Trans. Graph.* 37, 4, Article 164 (August 2018), 13 pages. <https://doi.org/10.1145/3197517.3201350>

## 1 INTRODUCTION

Reenactment approaches aim to transfer the motion of a source actor to an image or video of a target actor. Very recently, facial reenactment methods have been successfully employed to achieve highly-realistic manipulations of facial expressions based on commodity video data [Averbuch-Elor et al. 2017; Suwajanakorn et al. 2017; Thies et al. 2015, 2016, 2018; Vlasic et al. 2005]. Rather than animating a virtual, stylized avatar (e.g., as used in video games), these algorithms replace the face region of a person with a synthetic re-rendering, or modify the target image/video under guidance of a 3D face model. This enables changing the expression of a target person and creating a manipulated output video that suggests different content; e.g., a person who is sitting still could appear as if he/she is talking. Modern reenactment approaches achieve highly believable results, even in real-time, and have enjoyed wide media coverage due to the interest in general movie and video editing [Vlasic et al. 2005], teleconferencing [Thies et al. 2018], reactive profile pictures [Averbuch-Elor et al. 2017], or visual dubbing of foreign language movies [Garrido et al. 2015].

Even though current facial reenactment results are impressive, they are still fundamentally limited in the type of manipulations they enable. For instance, these approaches are only able to modify facial expressions, whereas the rigid pose of the head, including its orientation, remains unchanged and does not follow the input video. Thus, only subtle changes, such as opening the mouth or adding wrinkles on the forehead are realized, which severely limits



# Robust Solving of Optical Motion Capture Data by Denoising

DANIEL HOLDEN, Ubisoft La Forge, Ubisoft, Canada

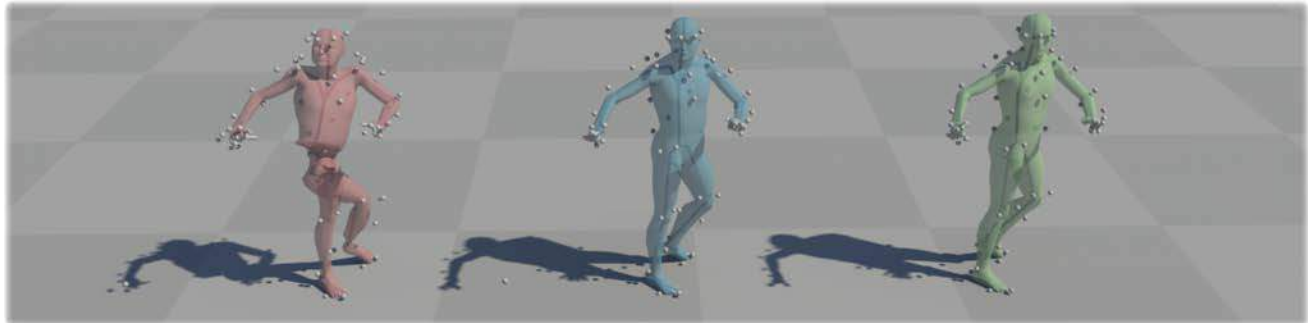


Fig. 1. Results of our method applied to raw motion capture data. Left: Raw uncleaned data. Middle: Our method. Right: Hand cleaned data.

Raw optical motion capture data often includes errors such as occluded markers, mislabeled markers, and high frequency noise or jitter. Typically these errors must be fixed by hand - an extremely time-consuming and tedious task. Due to this, there is a large demand for tools or techniques which can alleviate this burden. In this research we present a tool that sidesteps this problem, and produces joint transforms directly from raw marker data (a task commonly called “solving”) in a way that is extremely robust to errors in the input data using the machine learning technique of *denoising*. Starting with a set of marker configurations, and a large database of skeletal motion data such as the CMU motion capture database [CMU 2013b], we synthetically reconstruct marker locations using linear blend skinning and apply a unique *noise function* for corrupting this marker data - randomly removing and shifting markers to dynamically produce billions of examples of poses with errors similar to those found in real motion capture data. We then train a deep denoising feed-forward neural network to learn a mapping from this corrupted marker data to the corresponding transforms of the joints. Once trained, our neural network can be used as a replacement for the solving part of the motion capture pipeline, and, as it is very robust to errors, it completely removes the need for any manual clean-up of data. Our system is accurate enough to be used in production, generally achieving precision to within a few millimeters, while additionally being extremely fast to compute with low memory requirements.

CCS Concepts: • **Computing methodologies** → **Motion capture**;

Additional Key Words and Phrases: Optical Motion Capture, Motion Capture Cleanup, Motion Capture, Machine Learning, Neural Networks, Denoising

Author’s address: Daniel Holden, Ubisoft La Forge, Ubisoft, 5505 St Laurent Blvd, Montreal, QC, H2T 1S6, Canada, daniel.holden@ubisoft.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART1 \$15.00

<https://doi.org/10.1145/3197517.3201302>

## ACM Reference Format:

Daniel Holden. 2018. Robust Solving of Optical Motion Capture Data by Denoising. *ACM Trans. Graph.* 37, 4, Article 1 (August 2018), 12 pages. <https://doi.org/10.1145/3197517.3201302>

## 1 INTRODUCTION

Although a large number of different motion capture systems have been developed using many kinds of technologies, optical motion capture still remains the main technique used by most large movie and game studios due to its high accuracy, incredible flexibility, and comfort of use. Yet, optical motion capture has one major downside which severely limits the throughput of data that can be processed by an optical motion capture studio - namely that after each shoot the raw marker data requires “cleaning”, a manual process whereby a technician must fix by hand all of the errors found in the data such as occluded markers, mislabeled markers, and high frequency noise. Although this process can often be accelerated by commercial software which can provide powerful tools for aiding in the cleanup of motion capture data [Vicon 2018], it can still often take several hours per capture and is almost always the most expensive and time consuming part of the pipeline. Once the data has been cleaned, further automatic stages are required including “solving”, where rigid bodies are fitted to groups of markers, and “retargeting”, where inverse kinematics is used to recover the local joint transformations for a character with a given skeleton topology [Xiao et al. 2008].

In this paper we propose a data-driven approach to replace the solving stage of the optical motion capture pipeline. We train a deep denoising feed-forward neural network to map from marker positions to joint transforms directly. Our network is trained using a custom noise function designed to emulate typical errors that may appear in marker data including occluded markers, mislabeled markers, and marker noise. Unlike conventional methods which fit rigid bodies to subsets of markers, our technique is extremely robust to errors in the input data, which completely removes the need for any manual cleaning of marker data. This results in a motion capture pipeline which is completely automatic and can therefore achieve a much higher throughput of data than existing systems.

# Online Optical Marker-based Hand Tracking with Deep Labels

SHANGCHEN HAN, Facebook Reality Labs

BEIBEI LIU, Facebook Reality Labs

ROBERT WANG, Facebook Reality Labs

YUTING YE, Facebook Reality Labs

CHRISTOPHER D. TWIGG, Facebook Reality Labs

KENRICK KIN, Facebook Reality Labs



Fig. 1. **Snapshots of our online marker-based hand tracking system on sequences with two-handed and hand-object interactions.** We demonstrate a novel marker-labeling and tracking system that enables fully-automatic, real-time estimation of hand poses in challenging interaction scenarios with frequent occlusions. Markers labeled as left hand and right hand are rendered as orange and blue spheres respectively, while markers associated with predefined rigid bodies are rendered as green spheres.

Optical marker-based motion capture is the dominant way for obtaining high-fidelity human body animation for special effects, movies, and video games. However, motion capture has seen limited application to the human hand due to the difficulty of automatically identifying (or labeling) identical markers on self-similar fingers. We propose a technique that frames the labeling problem as a keypoint regression problem conducive to a solution using convolutional neural networks. We demonstrate robustness of our labeling solution to occlusion, ghost markers, hand shape, and even motions involving two hands or handheld objects. Our technique is equally applicable to sparse or dense marker sets and can run in real-time to support interaction prototyping with high-fidelity hand tracking and hand presence in virtual reality.

CCS Concepts: • **Computing methodologies** → **Modeling and simulation**; **Machine learning**;

Additional Key Words and Phrases: motion capture, hand tracking, marker labeling

Authors' addresses: Shangchen Han, Facebook Reality Labs, shangchen.han@oculus.com; Beibei Liu, Facebook Reality Labs, beibei.liu@oculus.com; Robert Wang, Facebook Reality Labs, rob.wang@oculus.com; Yuting Ye, Facebook Reality Labs, yuting.ye@oculus.com; Christopher D. Twigg, Facebook Reality Labs, chris.twigg@oculus.com; Kenrick Kin, Facebook Reality Labs, kenrick.kin@oculus.com.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Copyright held by the owner/author(s). Publication rights licensed to the Association for Computing Machinery.  
0730-0301/2018/8-ART1 \$15.00  
<https://doi.org/10.1145/3197517.3201399>

## ACM Reference Format:

Shangchen Han, Beibei Liu, Robert Wang, Yuting Ye, Christopher D. Twigg, and Kenrick Kin. 2018. Online Optical Marker-based Hand Tracking with Deep Labels. *ACM Trans. Graph.* 37, 4, Article 1 (August 2018), 10 pages. <https://doi.org/10.1145/3197517.3201399>

## 1 INTRODUCTION

Optical motion capture is the prevailing method for capturing high accuracy and high framerate motion. Applications across visual effects, gaming, VR/AR, usability research, and biomechanics are enabled by commercial systems that provide reliable sub-millimeter accurate tracking at up to 2000Hz. Notably, real-time full-body motion capture has enabled virtual production applications for games and movies [Antoniades 2016]. However, when it comes to detailed and dexterous finger motions, no commercial or academic software is able to produce comparable real-time tracking results, even in elaborate capturing environments.

The challenge of real-time capture lies in the identification or *labeling* of identical-appearing passive markers. While the labeling problem can be solved by active markers, their wired electronics or bulky sensors make them undesirable, and active systems are limited in the number of markers which can be tracked simultaneously. A common solution in full-body tracking for marker labeling is to start from a predefined T-pose or A-pose, then rely on high framerate tracking to propagate labels forward. This solution breaks down when tracking fails due to *occluded markers*, *ghost markers* from spurious reflections, or *dense marker sets* needed to capture subtle articulations. For body motion, tracking failures can be minimized by careful placement of cameras and marker layout design. Unfortunately, finger motions present additional challenges due to their high degree of articulation, self-similarity, and small scale. In

# ToonSynth: Example-Based Synthesis of Hand-Colored Cartoon Animations

MAREK DVOROŽŇÁK, Czech Technical University in Prague, Faculty of Electrical Engineering

WILMOT LI, Adobe Research

VLADIMIR G. KIM, Adobe Research

DANIEL SÝKORA, Czech Technical University in Prague, Faculty of Electrical Engineering

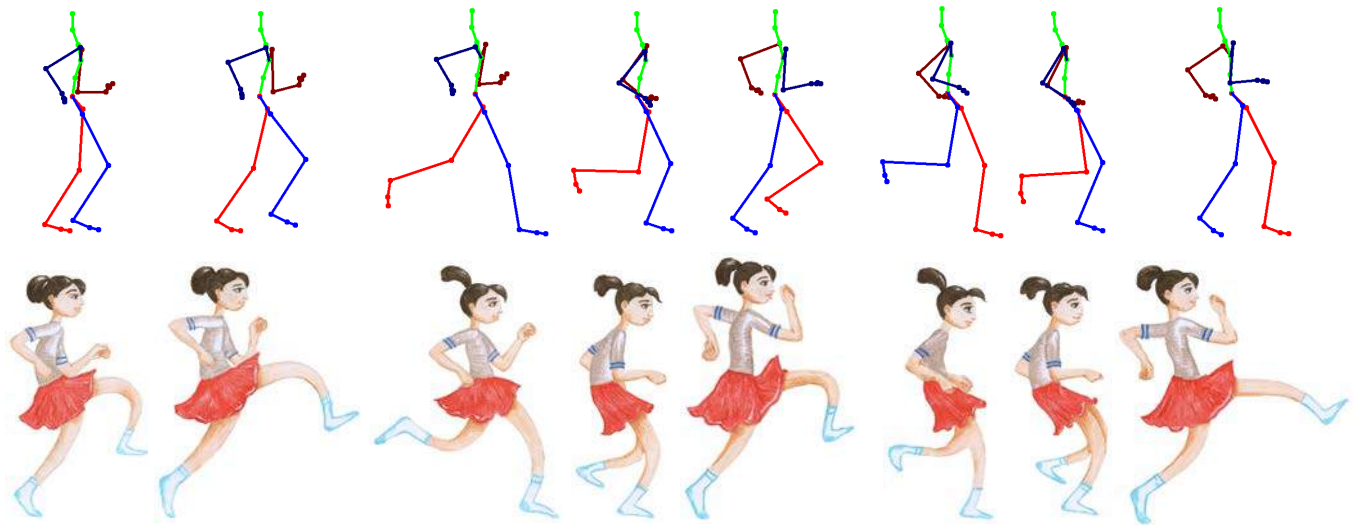


Fig. 1. An example of hand-colored animation synthesized using our approach (bottom row) following the user-specified skeletal animation (top row) and preserving the motion as well as appearance style prescribed by an artist (see a corresponding style exemplar in Fig. 10). Note how the synthesized images still resemble the hand-colored original.

We present a new example-based approach for synthesizing hand-colored cartoon animations. Our method produces results that preserve the specific visual appearance and stylized motion of manually authored animations without requiring artists to draw every frame from scratch. In our framework, the artist first stylizes a limited set of known source skeletal animations from which we extract a *style-aware puppet* that encodes the appearance and motion characteristics of the artwork. Given a new target skeletal motion, our method automatically transfers the style from the source examples to create a hand-colored target animation. Compared to previous work, our technique is the first to preserve both the detailed visual appearance and stylized motion of the original hand-drawn content. Our approach has

numerous practical applications including traditional animation production and content creation for games.

CCS Concepts: • **Computing methodologies** → **Motion processing**; **Image processing**;

Additional Key Words and Phrases: style transfer, skeletal animation

## ACM Reference Format:

Marek Dvorožňák, Wilmot Li, Vladimir G. Kim, and Daniel Sýkora. 2018. ToonSynth: Example-Based Synthesis of Hand-Colored Cartoon Animations. *ACM Trans. Graph.* 37, 4, Article 167 (August 2018), 11 pages. <https://doi.org/10.1145/3197517.3201326>

## 1 INTRODUCTION

While advances in computer graphics have contributed to the evolution of 3D animation as an expressive, mature medium, 2D animation remains an extremely popular and engaging way to tell stories. One common workflow for creating 2D animations is to decompose characters, objects and the background into separate layers that are transformed (either rigidly or non-rigidly) over time to produce the desired motion. A key advantage of this layer-based approach is that a single piece of artwork (i.e., layer) can be reused across many animated frames. As long as the appearance of the layer does not change dramatically (e.g., a character's torso turning from a front to side view), the artist does not need to redraw from

Authors' addresses: Marek Dvorožňák, Czech Technical University in Prague, Faculty of Electrical Engineering, [dvoromar@fel.cvut.cz](mailto:dvoromar@fel.cvut.cz); Wilmot Li, Adobe Research, [wilmotli@adobe.com](mailto:wilmotli@adobe.com); Vladimir G. Kim, Adobe Research, [vokim@adobe.com](mailto:vokim@adobe.com); Daniel Sýkora, Czech Technical University in Prague, Faculty of Electrical Engineering, [sykorad@fel.cvut.cz](mailto:sykorad@fel.cvut.cz).

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2018 Association for Computing Machinery.

0730-0301/2018/8-ART167 \$15.00

<https://doi.org/10.1145/3197517.3201326>