

Thinking FastAR: an augmented reality system for improving health decisions

Kenneth Friedman

May 6, 2016

When faced with a decision, it is often difficult for people to choose the best option for their long term well-being. Augmented reality can enable users to see more than what is physically present. Here, I propose and demonstrate an augmented reality system that helps users make healthier choices. Using object recognition as a heuristic for decision recognition, the system guides the user toward objects that align with a user's personalized health goals. The current implementation of this system involves predefined objects with hard coded health goals. For future work, recent advances in object recognition appear promising for a more universal version of the system, which will provide more flexibility and customization for users.

Introduction

HOW OFTEN HAVE YOU reached for a soda, even though you know a bottle of water would have been better for you? How many times have you taken the elevator to the third floor, instead of walking two flights of stairs? When making a decision, people are often near-sighted. It can be hard for people to make the best decision for their long term well-being, when it is so much easier to think about short term gains.

Recent work in cognitive psychology and behavioral economics¹ have catalogued the biases people face that affect their decision making abilities. The knowledge of these biases are beneficial in reflecting on previous decisions, but do little to improve real-time decision making. Real-time decision making relies almost entirely on what can be "seen" in the moment. The question then arises, can we build technology that enables us to "see" more?

Augmented reality (AR) is the perfect solution. AR can literally augment what the user can see. Here, I investigate whether AR can be used to augment decision making. I propose FastAR², a framework that can detect and provide feedback for certain types of decision opportunities in real time. I also present a successful, but early, implementation of the framework in the form of an iOS application.

Vision

¹ D. Kahneman. *Thinking, Fast and Slow*. 2011

² pronounced fast-ar or fast-a-r interchangeably

IN THEORY, people should make decisions that benefit them in the long term. However in reality, short term results are much easier to consider. This phenomenon in human behavior is due to a multitude of cognitive biases outlined in Professor Khaneman’s recent work. These biases can be summarized in Khaneman’s acronym WYSIATI: WHAT YOU SEE IS ALL THERE IS. That is, human decision making is based on information that is immediately available.

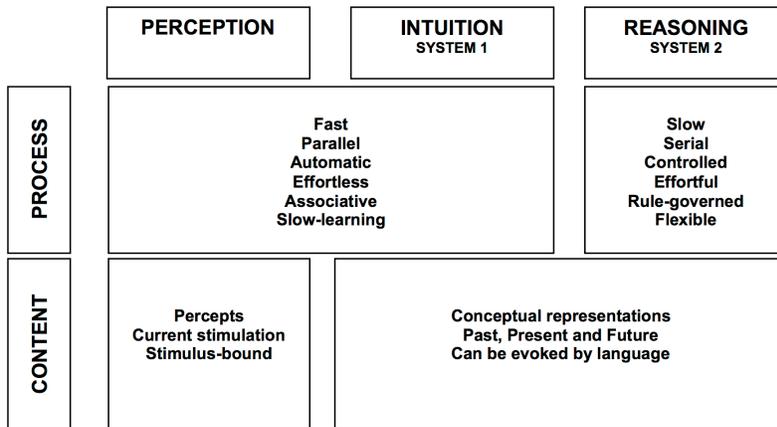


Figure 1: Kahneman’s model of cognitive thinking, in which System 1 acts quickly and System 2 plans more carefully. *FastAR* attempts to modify System 1 decision making without decreasing System 1’s speed.
(Figure from D. Kahneman)

I believe augmented reality can be used to augment this information, and therefore improve decision making. *FastAR* is a framework and implementation of this vision. The framework is a conceptualization of a computer system that can understand when a user is making a decision, and then present suggestions to the user to make a better long term decision. The implementation is an iOS application meant to demonstrate the fundamental ideas of the framework. Because the iOS app is an incomplete execution of the vision, the framework and implementation are discussed independently, below.

Background and Related Work

FASTAR DRAWS ON background work from both psychology and augmented reality. In psychology, the system draws on the work of Kahneman and Tversky’s nobel-prize winning work in judgement and decision making. Many of these ideas are summarized in Kahneman’s *Thinking, Fast & Slow*.

In a humorous account of a conversation with Kahneman, Richard Thaler³ explains that Kahneman was working on a book in 1996 and claimed that the book would be ready in six months. The book was published in 2000, 4 years later. Kahneman had fallen victim to the

³ R.H. Thaler and C.R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. 2008

planning fallacy⁴ (a phrase that he himself had coined). If Kahneman can't avoid cognitive biases, no one can. It's clear from Kahneman's work, knowledge alone isn't enough to overcome cognitive biases.

In augmented reality, FastAR draws on a number of previously developed systems. Most notably, many elements of AfterMath⁵. AfterMath is an AR system that enables users to see into the future. While AfterMath demonstrates the effects of a user's action on the outside world, FastAR attempts show the personal effects of a user's decision.

FastAR Framework

THERE ARE THREE MAIN COMPONENTS to the FastAR framework. First, there is a backend representation of the user. This is useful for personalizing the system to a given user. Second, FastAR has a universal object recognition system. This can detect an object in sight, and check to see if that object is relevant to the user. Third, the AR component presents the user with a suggestion to guide them toward a better solution. Each component will be discussed in detail.

First, an internal representation of the user is a key component of the system. Individuals differ in their health goals and health desires. It would be problematic for a system to dictate what choices are best for the user. Therefore, upon acquisition of the system, the user is prompted to select health goals, such as "exercise more" or "drink less soda." It is also possible for user's to create customized goals by informing the system of categories of objects that would be either beneficial or harmful to the goal. For example, a particular user might want to reduce the number of times they eat at fast food restaurants. This would be a user defined goal, so the user simply informs the system that "McDonalds" and "Burger King" are "negative objects" (explained below). With this simple rule-based method, FastAR can easily decide if a particular object helps or hinders the goals of the user.

The second component, universal object recognition, enables the system to understand the world around it. Recent breakthroughs in deep learning systems allow for real-time object recognition (such as UC Berkley's accurate object recognition system⁶). Once an object is recognized, it determines whether an object is a "positive" or "negative" object based on whether the object is or is not beneficial for the given user's health goals.

The third component provides the user with feedback. Positive objects result in a reward for the user, while negative objects prompt suggestions towards alternatives. The feedback starts at the sublim-

⁴ Wikipedia: Planning fallacy. URL https://en.wikipedia.org/wiki/Planning_fallacy

⁵ Sang-won Leigh and Pattie Maes. Aftermath: Visualizing consequences of actions through augmented reality. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 941–946. ACM, 2015

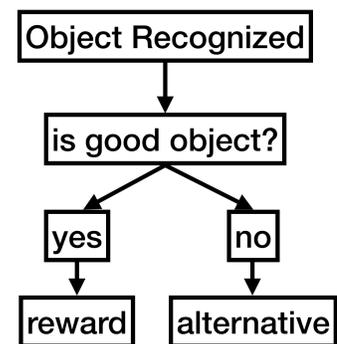


Figure 2: The program flow of FastAR. Once an object has been detected, the system checks whether or not the object is beneficial for the user. If the object is, the user is rewarded. If the object is not, an alternative object is suggested.

⁶ Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014

inal level. For example, a negative object might result in a quiet but irritating sound effect, or a blurring of the visual field over the object. The suggestions increase in force over time. For example, after many failed attempts at avoiding the soda, a user might be shown a video describing the negative side effects of a high-sugar drinks.

Together, these three components enable the system to understand the goals of the user, recognize objects, and suggest alternatives to a given object based on the user's goal.

FastAR Implementation

I HAVE IMPLEMENTED a successful but early version of FastAR as an iOS application. While AR goggles would improve the just-in-time nature of the system, the hand-held version demonstrates the system's potential.

The internal representation of the current system exists, but is limited to hard coded goals. These goals can be turned on or off, but new goals can not currently be added without modifying the code. As a demonstration, the system understands the goals "exercising more" (which understands stairs as a positive object and elevator as a negative object) and "drinking less soda" (which recognizes Mtn Dew and Coke as negative object and water as a positive object).

The object recognition module is powered by Vuforia⁷, which uses targets to recognize objects. This is perfect for objects with brand logos, such as sodas. However the method breaks down when detecting general objects. Therefore, the elevator demonstration will only detect the specific elevator at the MIT Media Lab. Deep learning systems would make this far more generalizable, and they can already detect objects in real-time locally (such as Jetpac's DeepBelief System⁸).

For the AR content, the app displays a video based on the selected object. In the demonstration, FastAR first detects a soda and then plays a video explaining the negative effects of soda. This likely an extreme augmentation, and it would be rarely used in a complete version of the framework.

Usage Scenarios

A COMPLETE IMPLEMENTATION of the framework would be widely general and extremely flexible. There are only two requirements for a given use case. First, the decision must be based on a visual, physical object. For example, choosing a drink is based on a physical object:

⁷ Vuforia platform. URL <http://www.vuforia.com>

⁸ Jetpac's deepbeliefsdk. URL <https://github.com/jetpacapp/DeepBeliefSDK>

the drink itself. The system would not be able to help a user decide to stand up and take a break from work (which other systems can easily do). Second, the object must be widely understood to be either positive or negative for your health. Presented with an unknown object, the system would not be able to provide feedback.

The FastAR implementation demo is limited to hard coded demos. Soda v. Water and Elevator v. Stairs are the two primary examples described throughout this report and easily accessible on the iOS app.

Future Work

OBVIOUS FUTURE IMPROVEMENTS involve upgrading the iOS app to implement the complete framework. This includes transitioning from the Vuforia SDK to a deep neural network system. It would also be necessary to upgrade the back-end representation of the user to be more flexible.

More interestingly, the best way to display the suggestions should be researched. Currently, the system presents obtrusive videos, however more nuanced approaches might be much more successful. Subliminal interfaces, or barely notable FOV modifications might change user behavior without conscious understanding of the change.

Finally, the framework add the feature of being able to see into the future further. It would be interesting to experiment with being able to simulate what happens to the user when a decision is made. For example, unhealthy decisions could result in a change to physical appearance.

Contributions

I HAVE PRESENTED FastAR, a framework for using augmented reality to enhance a user's decision making capabilities. I have drawn on research from psychology and prior AR concepts to develop a system that helps users make healthier choices. Also, I have implemented a prototype of this framework on an iOS device. This handheld version provides a glimpse at what is possible with modern technology and demonstrates the advantages of this framework.

References

Wikipedia: Planning fallacy. URL https://en.wikipedia.org/wiki/Planning_fallacy.

Jetpac's deepbeliefsdk. URL <https://github.com/jetpacapp/DeepBeliefSDK>.

Vuforia platform. URL <http://www.vuforia.com>.

Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 580–587, 2014.

D. Kahneman. *Thinking, Fast and Slow*. 2011.

Sang-won Leigh and Pattie Maes. Aftermath: Visualizing consequences of actions through augmented reality. In *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pages 941–946. ACM, 2015.

R.H. Thaler and C.R. Sunstein. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. 2008.