Beyond the Cup: Tabletop Augmented Reality for Coffee Awareness

Gloria Chua
Stanford University
Stanford, CA 94305, USA
cjmg@stanford.edu

Victoria Wee
Stanford University
Stanford, CA 94305, USA
vwee@stanford.edu

Sage Isabella
Stanford University
Stanford, CA 94305, USA
sageis@stanford.edu

Monica Yupa
Stanford University
Stanford, CA 94305, USA
myupa@stanford.edu

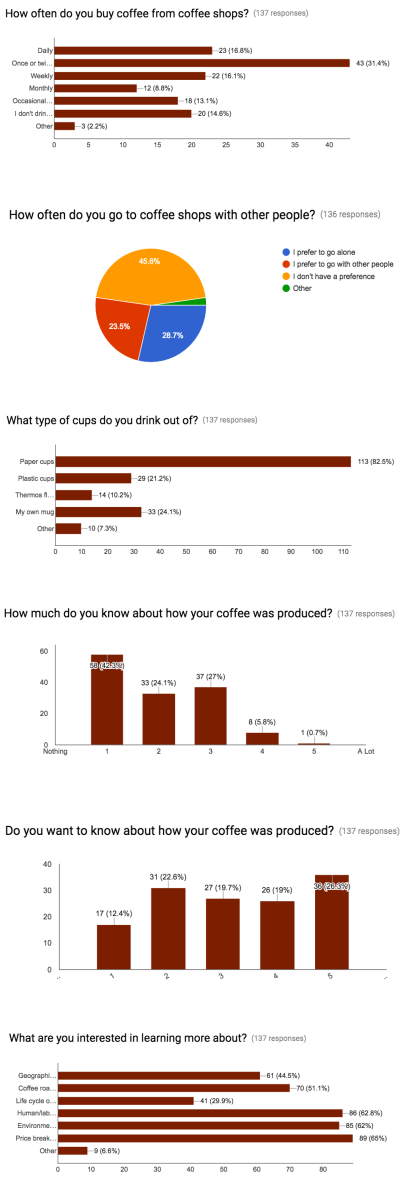
Abstract
UPDATED—June 7, 2016. This paper describes the design, development and implementation of Beyond the Cup (BTC), an interactive tabletop augmented reality system that increases user empathy surrounding the coffee industry. BTC is instrumented with two RGB cameras, 2 optical projectors, and two projection screens placed perpendicular to each other, and is designed for the use case of an interactive exhibit in a coffee shop or a coffee-themed museum. We illustrate the process of need-finding, prototyping, testing and implementing this application, and discuss the strengths and limitations of this final application.

Author Keywords
Tabletop augmented reality; projector-camera system; interactive tabletops; tangible augmented reality

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Multimedia information systems: Artificial, augmented and virtual realities

Introduction
Rendering and interacting with real world and virtual world objects presented within a single display is a fundamental idea in mixed reality (MR) [1]. Interacting with these objects on a physical tabletop environment is not a novel concept, but there has been little exploration done in the retail domain. In this paper, we explore a novel tabletop augmented reality application with a particular use case in the retail space: a 2-minute interactive application that encourages users to understand the ethical and social dimensions of an

Figure 1: Survey Results



ordinary everyday product - a cup of coffee. We envision this application to be used in context (e.g. placed in a coffee shop for users to interact with while waiting for their cup of coffee to arrive).

Motivation and contribution

The creation of BTC was primarily motivated by three separate objectives. First, we were interested to explore the storytelling affordances of augmented reality, particularly for tabletop and tangible augmented reality. Azuma's lecture on storytelling in augmented reality [2] was particularly influential in shaping this motivation. Second, we found the use case of unravelling a physical object's history and story a compelling use case of merging physical and virtual worlds. Third, we were motivated to develop an application that had the potential to catalyze positive social impact and behavioral change.

In searching for particular domain areas to explore these motivations, we settled on the retail space, given that storytelling is a commonly used technique in retail, the retail industry is comprised of physical objects that each have their own "history", and existing retail applications are primarily motivated by profit instead of social impact. The lack of non-profit-driven applications presents an excellent opportunity to create a novel AR application within the retail space.

Related Work

Tangible Augmented Reality

We draw inspiration from past work on tabletop tangible interfaces feature tangible objects that can serve as both input and output devices, and can be augmented with visual content corresponding to its shape, position and orientation. With digital content being projected directly on physical objects, this removes the need for a mediated device (e.g. heads-up display/mobile device) to add extra layers of information. The tabletop form also provides social affordances that support collaboration and interpersonal interaction [3]. The compromise is flexibility - tabletop interfaces require carefully calibrated projectors and is more suitable for exhibition settings. However, given that our application is meant to stay in a particular location (e.g. retail space), we are able to work within this constraint.

Another interesting case study is Hiroshi Ishii's lesser-known work in tribute to Mark Weiser - his development of transparent interfaces, using bottles as the primary physical object. [4] In it, Ishii highlights the bottle interface is a pursuit of a minimal interface based on a well-understood generic object. There is a notable parallel between cups (our interface object) and bottles, for both are richly-afforded volume-holding devices. The framework of augmented information that Ishii takes is particularly relevant to guiding our design work - digital content, tangible interfaces using bottles, and sensing technologies.

Tabletop Augmented Reality

A set of tabletop AR systems utilize a projector-translucent table system to augment physical objects with virtual information. In the ReacTable system, physical objects with identification markers serve as input devices and are augmented by the 2-D tabletop projection with visual cues about their semantic relationships. In the Tangible Urban Planning/ Projected Play system, the objects themselves serve as projection surfaces [5]. Object tracking is done using fiducial tracking by cameras placed under the translucent table,

and certain objects are used as input devices (e.g. using a cylinder as a camera controller), while others are used as output devices (e.g. Lego brick buildings) on which visual information is displayed.

We are particularly inspired by another set of tabletop AR systems that support freehand interaction with 3-D virtual objects, instead of with digitally-augmented physical ones. An example is MirageTable, which uses a curved projection surface, depth cameras (e.g. Kinect), head tracking and projective texturing to provide correct 3-D perspective views [6]. Another application, LightSpace, provides a detailed and robust description of projector-depth camera setups that we are likely to be able to use in our application, and other interesting interactions e.g. spatial menus and thought-body transitions [7].

Retail Applications of Augmented Reality

There exists multiple applications meant to enhance a customer's retail experience in-store and online. Potential applications include personalized in-store e-commerce through dynamic contextualization [8], displaying item image in contextual environments [9], and augmented reality mirrors for virtual trying-on of apparel [10]. As mentioned, many of these applications are primarily driven to increase retail sales and profits. This presents a unique opportunity to design an application contrary to that very purpose - one that is more educational in nature, to inform consumer choice through educating users about the ethical and social dimensions behind a particular product.

Methodology

Initial need-finding survey

We began by narrowing down the retail industry to a specific sub-industry - coffee retail. This focus was primarily motivated by the ubiquity of coffee purchases among urban consumers, our target audience. The presence of an existing coffee shop on campus (Starbucks) also would enable more authentic field need-finding and provide a meaningful context for the project.

After picking the coffee industry, we sought to understand coffee drinking habits and potential topics for exploration via survey distribution. We chose a survey instead of deeper contextual inquiry techniques to gather a large amount of information in a short amount of time. The survey was made using Google Forms and reached 137 individuals, 79 self identified females and 50 males (one respondent chose not to answer), across a diverse age range (18-52 y/o). 90% of participants were Stanford students and faculty living in the Bay Area, and the remaining lived in Singapore. Given that both populations are relatively urban, and that we had sufficiently many participants, we believe the survey results to be represent the opinions of our target audience.

We found that of that over 75% of respondents either preferred going to coffee shops alone or had no preference. This information shows that the interaction should be able to be done with one individual. We found that 83% of the respondents use paper cups, 83% frequent coffee shops at least weekly, and 85% feel like they know very little about how their coffee was produced. Approximately 46% of respondents felt like they would want to know more about how their coffee

was produced. If our team is to reach a large audience, the interactions will have to be engaging and make people want to engage in coffee industry issues.

The survey also questioned users on what specific dimensions of the ethical and social dimensions of coffee they would be interested in learning. The top three dimensions were: the price breakdown of coffee (65%), 2. human and labor rights (62.8%), and 3. environmental sustainability (62%). The next most popular dimension had 16% lower number of votes - significant enough for us to focus directly on creating compelling interactions for these three aspects.

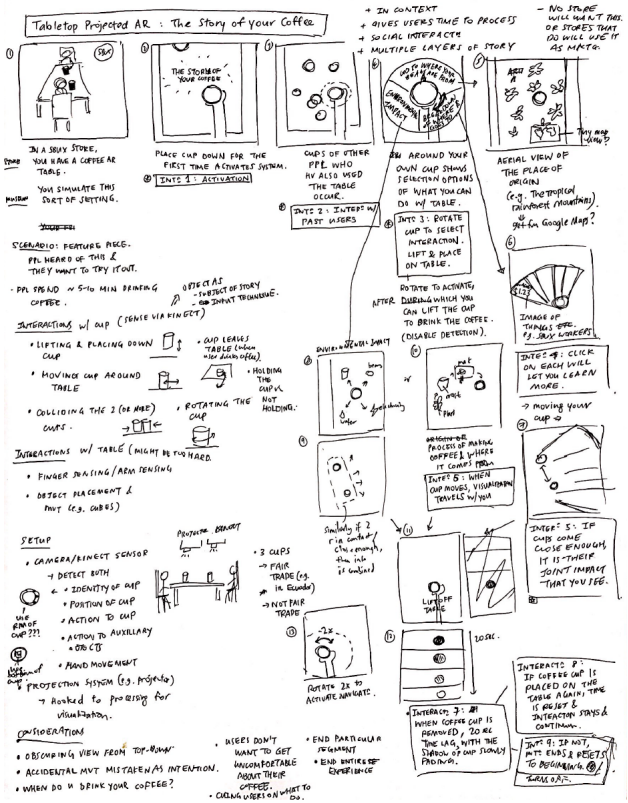
Contextual observations

We also observed our target users in context (i.e. at a coffee shop - Starbucks at Stanford's Tresidder Union), to understand how this particular application can fit into the context of a coffee purchase experience in-store. We noted that there were 2 particular intervals where users spent simply waiting, either in line to order or for their coffee to arrive. In both intervals, users were observed to wait for around 2-3 minutes, and often spent their time using their phones, even if they came with a partner. Additionally, users tended to cluster in one corner near the milk/sugar counter in the second interval - an ideal location for a tabletop AR interaction. From these observations, we identified the post-ordering wait time to be a suitable use case/context for our particular observation, and determined that our interaction should be no more than 2 minutes long.

Initial prototype

Based on our learning from the survey, observations and topics discussed in class, we designed early

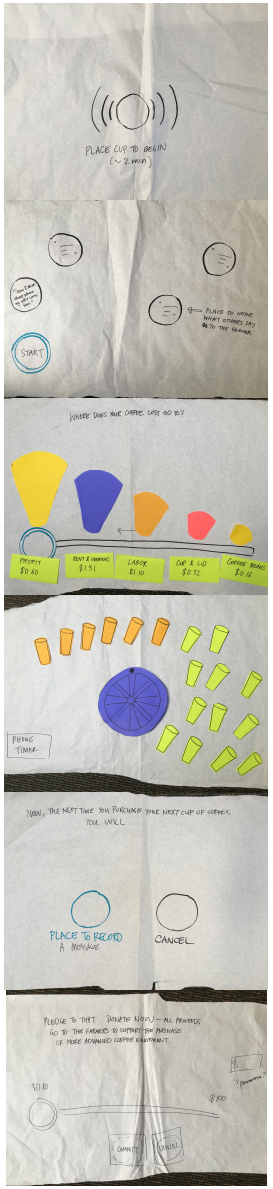
Figure 2: Initial Storyboard



concepts for an interactive tabletop AR system. The design principles that guided our concepts were:

- Tell a story (allowing for a non-linear narrative)
- Let coffee cup be connection between physical and virtual world

Figure 3: Paper Prototype



- Build in bodily interactions
- Don't let social message be too overt (it may overwhelm the user)
- Allow for social interaction with other users
- Favor feasibility over style (make sure we build something that works)

We devised 6 main concepts/interactions after initial explorations, listed here chronologically, and rapidly prototyped them at scale on paper. Refer to Figure [3] for images of the paper prototypes.

- *Social* - Users are able to listen to messages left by past users, to motivate and frame the experience they are about to go through
- *Price line* - Users interact with a price line that shows the breakdown of costs that go into a cup of coffee, augmented with zoom-in video footage of each cost component, and interacted with by moving the coffee cup along the price line
- *Farmer introduction* - Users are transported to a coffee plantation and are introduced to a coffee farmer, who shares about how coffee is made
- *Hulling activity* - Users empathize with coffee farmer through "hulling" beans by rotating their coffee cup on the tabletop. They are then showed how much more work one has to take to be able to purchase the cup of coffee they purchased.
- *Leave a message* - Users can leave a message for the coffee farmer (and other future users) based on what they learnt, internalizing the message by reflecting on it
- *Donate* - Users can take action by donating to the coffee farmer an amount in the order of magnitude of the price of their cup of coffee

Description of User Study

To understand which concepts were most compelling, we set up a user study with 5 participants (2 Female, 3 Male). Each participant was explained the purpose of

Figure 4: User Study

Figure 5: Set-Up

the study and encouraged to think aloud throughout the 15-min study. Participants were shown the start screen and were told to navigate through the application, and difficulties and delight points were noted. After the activity, participants were asked a standardized series of questions to understand their key takeaways, and their thoughts about each specific concept.

Learnings of User Study

All 5 participants were able to articulate the underlying message of the application, even without us explaining to them beforehand. It was clear that the hulling activity and price line were the 2 most compelling features - when asked for the most memorable feature of the application, 3 voted for the hulling activity, and 2 voted for the price line.

In the price line, all participants easily grasped the interaction of moving a coffee cup across the table to dissect the price breakdown. Users also enjoyed getting to see videos and images of each component (e.g. picture of a Starbucks shop when the cup is at the 'Rent' component).

In the hulling activity, users particularly enjoyed the bodily action it entailed. Despite the fact that rotating a coffee cup is not a labor-intensive action, the act of moving itself was effective in conveying the message that making coffee was a labor-intensive process.

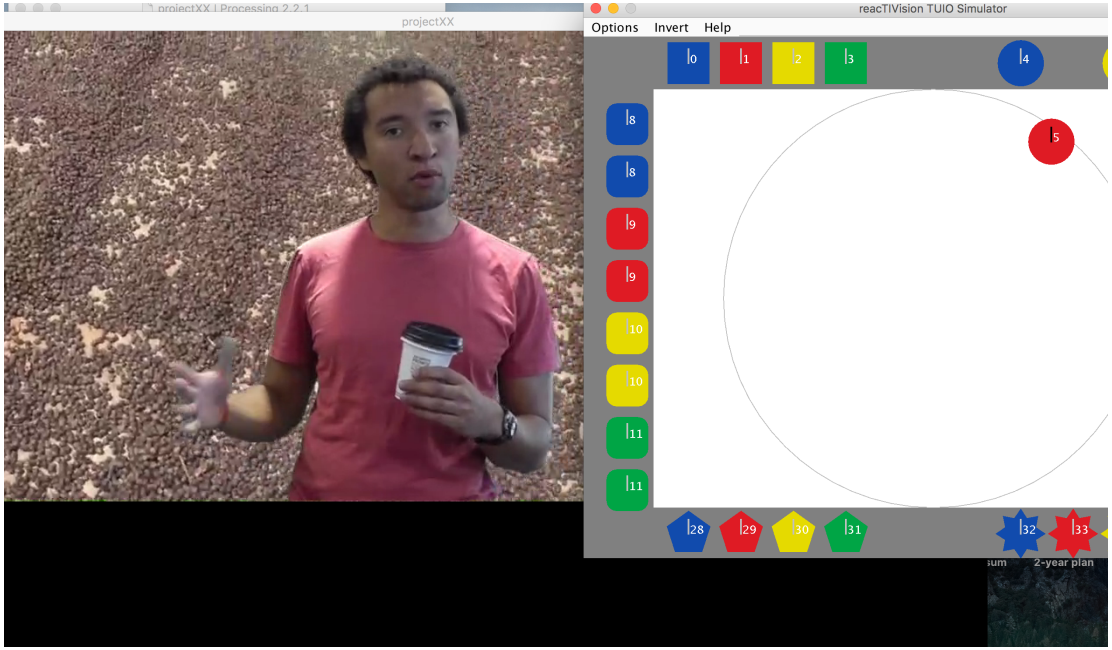
One user also highlighted his surprise at the Farmer introduction component, and explained that he felt connected to the farmer, and that "he almost wanted to ask the farmer more questions". While we were initially worried that the farmer's monologue was too long in duration, none of the participants brought it up as a pain point.

The social component of listening to other people's messages were fun for 3 participants, but when told to leave a message for others, most did not know what type of message to leave. Given that the social component requires dynamic content creation by participants, and that participants were unmotivated to create such content, we deemed this component as ineffective.

Despite designing the "Leave a Message" and "Donate" interactions as methods for users to take action, 4 out of 5 participants described a sense of helplessness at the end of the interaction. Many felt burdened by their new knowledge, but did not feel that leaving a message for other users, or donating a tiny amount was an appropriate way to take action.

In all, the pedagogical objective of the application has been effectively achieved, primarily due to the price line and hulling activity. The empathy-building aspect of the application was successfully accomplished by the farmer introduction feature. However, the call-to-action features (donation, leave a message) were not effective. More needs to be brainstormed in terms of creating a stronger call to action, but that is beyond the scope of this project and paper. We therefore decided to focus on building 3 main features of the application: the price line, farmer introduction, and hulling activity.

Figure 6: Development Environment



System Implementation

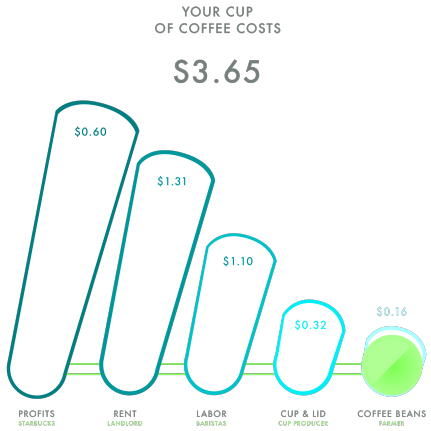
We used a 2-surface 2-projector 2-webcam configuration to set up our projected tabletop AR system. The 2 surfaces are perpendicularly placed to each other to simulate a tabletop and front screen. A schematic of our set-up is found in Figure [5].

Our system was developed with Processing (v.2.2.1) and using TUIO, an open source framework designed specifically for tabletop tangible user interfaces. Our system also leveraged reacTIVision, a computer vision framework for object tracking and basic multi-touch. We also used TUIO Simulator to test the system's fiducial interactions. Note that we used an older version of Processing because the latest version (v.3.1.1) was not compatible with reacTIVision.

Each webcam is connected to a laptop and projector. The first webcam-laptop-projector unit runs a program that projects onto the tabletop surface. The second webcam-laptop projector unit runs a program that projects on a vertical surface immediately adjacent to the tabletop surface. The two units work in sync to generate a more immersive user experience with visual feedback on the two surfaces.

Both units are synchronized by and respond to the webcams' detection of the user's interaction with the tabletop projection. Since both webcams watch the same surface, the two units are trivially synchronized as long as each projector-webcam unit is correctly calibrated to recognize and track the same fiducial in the system.

Figure 7: Price Line



The user interacts with the tabletop projection using a coffee cup as gestural input. The system tracks one fiducial marker on the lid of the coffee cup. When the marker triggers specific areas of the projection, different parts of the program are projected in response. We hardcoded the dimensions of the tabletop surface and used simple calculations relative to these dimensions to determine when certain areas of the surface projection were triggered by the fiducial.

Price Line

When the cup is on the timeline, the tabletop screen displays the price breakdown in the form of an augmented shadow that moves together with the cup. Users are able to slide their cup on the projected tabletop, guided by a line. At particular spatial positions on the price line, the data on how much a particular component of the coffee (e.g. the beans, shop rent, labor costs etc.) is displayed below the cup. Simultaneously, the screen displays video footage of that particular component (e.g. of an extra Starbucks shop, corresponding to the Rent component) to provide the context for that particular component. This particular feature interaction ends when the users place the cup at the end of the price line, during which they will see a graphical breakdown of each price component (Figure [7]). Data for the price breakdown of coffee is found from [11].

Farmer Introduction

The price line feature transitions to the farmer introduction one through the coffee farmer virtually grabbing the coffee cup from the user. We could not film on-site in an actual coffee farm, so we recruited an actor to play the coffee farmer and extracted video clips on the coffee farming process (e.g. plucking the beans,

hulling) from YouTube, using those to simulate being on a coffee farm.

The coffee farmer introduces himself, directly addressing the user, and shares about where he comes from, and the various steps in the coffee-farming process. Figure [[b]] shows photos of a user interacting with this particular segment.

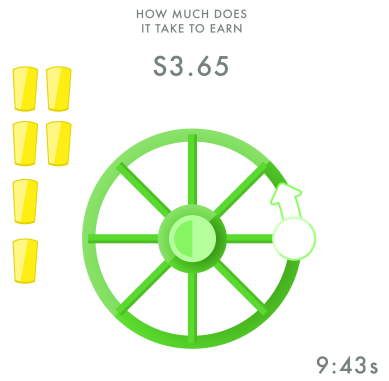
Figure 8: Farmer Interaction

While the coffee-making process was not directly listed as one of the top 3 things our initial need-finding pool wanted to learn about, we found it an effective way to convey (through image) the human/labor rights aspect by portraying the farmer himself. Additionally, this segment serves an empathy-building function by providing a direct connection between the actual coffee farmer and the user, who in their purchase of coffee often does not consider the coffee farmer himself.

Hulling Activity

After the farmer's introduction, a hulling wheel appears on the tabletop screen, and users are encouraged to use their cup to virtually hull the coffee beans, to see how many cups of beans they might need to produce to be able to afford the cup of coffee they are drinking. (Figure [9]) The coffee cup serves as an input mechanism – users hull the beans by rotating the cup around the table. After a designated amount of time (10 seconds), the coffee farmer explains and shows the number of cups produced by the user, and the number of additional cups necessary.

Figure 9: Hulling Graphics



This segment is intended to evoke a strong sense of empathy with the coffee farmer, alluding to the point about human/labor rights, illustrating how time and labor-intensive it is to produce coffee, and how little profit coffee farmers actually make.

Discussion

Strengths and Limitations

BTC is successful in that it is a working interactive table top system that takes advantage of augmented reality. for empathy building. This product allows users to take a relevant physical object, their coffee cup, and connect it to an experience at a coffee shop. Through this product they can more deeply augment their reality by learning from a farmer how their coffee was made. Lastly, their "real" coffee cup is transformed into a handle of a huller, completing the cycle of contextualizing their coffee in the real world.

BTC is also a thoughtfully prototyped product that strived to fit user needs and morph to user feedback. We were able to decipher early on topics that coffee drinkers were interested in learning more about and determining interactions that people enjoyed.

In terms of limitations, the prototype is not as sleek as it could be. It is currently an exposed mix of wood, foam core, 60% acrylic plastic, with two projectors and two RGB cameras, and takes up 3' X 6' X 4'.

There are also shadows on the projections which distract from the effect of the table top, but also limit the use of the RGB cameras to pick up important color data. Other limitations include shortage of testing on the final prototype. There are likely some interactions that could be helpful to users that do not exist in the current prototype or current implementations that don't make sense.

Learnings

This project was useful in learning within the scope of augmented reality and beyond. The team learned skills in rapid prototyping, green screening, projection, programming in Processing and ReacTIVision with TUIO. Some specific realizations are listed below.

Allowing users to take action has unique challenges

- How do you allow someone to do something that they want to do and not have to do?
- How to make the interaction comfortable and natural?
- Ex: Providing users the opportunity to donate: some users felt was nice others felt like this was unnecessary
- Ex: Providing users the opportunity to share a message: high pressure because would be shared with future users, prompt was not specific enough to answer, users felt uncomfortable speaking allow

Light occlusion is possible

- We found that the projectors light white washed the fiducial coloring which made it more difficult for the RBG cameras to recognize it.
- We had to keep darkening the fiducial with Sharpie

Software dependencies are a roadblock for new technologies

- We originally planned to use the Kinect with object direction to track the coffee cup, but Kinect did not have capability
- Switched to 3rd Party library for object color detection, but could not find a computer system to support it
- Finally shifted to Processing and ReacTIVision

- There was a shortage of documentation due to new technology

Contribution and Future Work

Although Microsoft has already explored the possibilities of tabletop augmented reality, our application extends into social change retail, an area that AR research usually does not touch. Using tabletop reality as an empathy building platform is largely unexplored. Our tabletop also has two perpendicular surfaces which has a more separated from reality feel than just a flat, parallel to the ground, surface.

This iteration of the project is proof of concept. Future work includes testing on-site, doing 3-D graphics, and either making a sleeker and more easily installable system with screens to limit occlusion or extending beyond tabletop AR and perhaps imagining this for a glasses-based system.

Conclusion

Beyond the Cup is an augmented reality table top system that allows people to learn about where their coffee comes from and to gain empathy from both a physical and virtual experience. This project's design is rooted in need-finding and adhering to user feedback.

After several rounds of iterations, our team settled on a system that allows users to find out how the price of their coffee was distributed and then to meet a farmer to learn how their coffee is made and farmer wages.

Through trial and error several lessons were learned about user action, software dependencies, light occlusion from light making fiducials unrecognizable. Our team added to the field of augmented reality by creating an empathy building device to connect the act of buying a cup of coffee to the story of the coffee itself.

The next steps for this project would be a bit more time with the code to clean up transitions and make interactions a little smoother. Then the device could go into a more thorough testing round with more users to see what could be added or removed from the experience. Lastly, the experience could be rebuilt for headsets or an improved physical structure could be implemented, perhaps with screens instead of projectors.

References

1. Milgram, Paul, et al. "Augmented reality: A class of displays on the reality-virtuality continuum." Photonics for industrial applications. International Society for Optics and Photonics, 1995.
2. Azuma, Ronald. Chapter 11 in 2nd Edition of Fundamentals of Wearable Computers and Augmented Reality, Woodrow Barfield, (editor), CRC Press, August 2015, pp. 259-276.
3. Jordà, Sergi, et al. "The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces." Proceedings of the 1st international conference on Tangible and embedded interaction. ACM, 2007.
4. Ishii, Hiroshi, Ali Mazalek, and Jay Lee. "Bottles as a minimal interface to access digital information." CHI'01 extended abstracts on Human factors in computing systems. ACM, 2001.
5. Dalsgaard, Peter, and Kim Halskov. "Tangible 3D tabletops: combining tangible tabletop interaction and 3D projection." Proceedings of the 7th Nordic

Conference on Human-Computer Interaction: Making Sense Through Design. ACM, 2012.

6. Benko, Hrvoje, Ricardo Jota, and Andrew Wilson. "MirageTable: freehand interaction on a projected augmented reality tabletop." Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 2012.
7. Wilson, Andrew D., and Hrvoje Benko. "Combining multiple depth cameras and projectors for interactions on, above and between surfaces."Proceedings of the 23nd annual ACM symposium on User interface software and technology. ACM, 2010.
8. Zhu, Wei, et al. "Personalized in-store e-commerce with the promopad: an augmented reality shopping assistant." Electronic Journal for E-commerce Tools and Applications 1.3 (2004): 1-19.
9. Tapley, John, and Eric J. Farraro. "Augmented reality system method and appartus for displaying an item image in acontextual environment." U.S. Patent Application No. 14/868,105.
10. Sukaviriya, Noi, et al. "Augmenting a retail environment using steerable interactive displays." CHI'03 Extended Abstracts on Human Factors in Computing Systems. ACM, 2003.
11. Bally, Paul. "The Economics of a Starbucks Cappuccino Grande." Web log post. Coffee Makers USA. N.p., 13 May 2014. Web. <http://coffeemakersusa.com/pricing-breakdown-cup-coffee/>.

Appendix 1 User Testing Guide

Metrics to Track

- Duration of experience
- Clarity and usability
- Achievement of learning outcome
- Emotional resonance
- Effectiveness compared to a non-AR experience

Roles

- Demo facilitator
- Demo actor/screen changer
- Timer / notetaker

Procedure

Get students/passers-by to come and test

We're a team of Stanford students testing out an augmented reality application surrounding knowing the story behind your coffee. Will you be willing to spend 10-15 minutes to test this out with us, and answer a few questions? We are happy to get you a drink in return.

Before we begin, we'd like to ask you some questions just to get your background. There are no right or wrong answers!

What is your name?

What were you doing in Tresidder?

How old are you?

What is your gender?

How often do you drink coffee?

What are 3 adjectives to describe how you are feeling at the moment.

Test

Facilitator introduces the activity with script below as actor sets up

This is purely a conceptual prototype and is in paper. Everything that you see, besides the cup you are holding, is digitally projected. Please speak your thoughts aloud, including any confusion, delight etc.. I will not be able to answer any of your questions, as we want to see how you work through this system! At the end of the activity we will be asking you a few questions about it. Sounds good?

Walk through test, timer starts timing

Facilitator take notes

Timer stop timing when experience ends

Question/data gathering phase

(Time spent: _____)

Facilitator: Now we'd like to ask you some questions about your experience. Remember, there are no right or wrong answers, we're just trying to understand more about your experience! Timer takes notes

How did the experience make you feel?

(If vague responses, give me 3 adjectives that describe how you feel at the moment)

What was your key takeaway from this experience?

Can you describe what happened during this (point to the timeline) phase of this experience? Repeat for (timeline, cups, social)

"I like, I wish" for each segment

Which part was the most compelling? Least compelling? Why?

In your mind, where do you see such an experience taking place?

How would you compare this to your previous experiences learning about the background of merchandise?

Overall, how would you describe your experience going through this demo? Would you recommend this to a friend?

Thank and follow up

Thank you so much for your participation! Feel free to go and get a drink and we can Venmo you when you bring back the receipt.