A Framework for Designing Intelligent Task-Oriented Augmented Reality User Interfaces

Leonardo Bonanni
MIT Media Laboratory
20 Ames Street
Cambridge, MA (USA)
+1 617 253 4564
amerigo@media.mit.edu

Chia-Hsun Lee
MIT Media Laboratory
20 Ames Street
Cambridge, MA (USA)
+1 617 253 4564
jackylee@media.mit.edu

Ted Selker
MIT Media Laboratory
20 Ames Street
Cambridge, MA (USA)
+1 617 253 6968
selker@media.mit.edu

ABSTRACT
A task-oriented space can benefit from an augmented reality interface that layers the existing tools and surfaces with useful information to make cooking more easy, safe and efficient. To serve experienced users as well as novices, augmented reality interfaces need to adapt modalities to the user's expertise and allow for multiple ways to perform tasks. We present a framework for designing an intelligent user interface that informs and choreographs multiple tasks in a single space according to a model of tasks and users. A residential kitchen has been outfitted with systems to gather data from tools and surfaces and project multi-modal interfaces back onto the tools and surfaces themselves. Based on user evaluations of this augmented reality kitchen, we propose a system to tailor information modalities based on the spatial and temporal qualities of the task, and the expertise, location and progress of the user. The intelligent augmented reality user interface choreographs multiple tasks in the same space at the same time.

Categories and Subject Descriptors
H.5.1 [**Information Systems**]: Multimedia Information Systems – *Artificial, augmented, and virtual realities.* H.5.2 User Interfaces – *Graphical user interfaces*.

General Terms
Performance, Design, Experimentation, Human Factors, Theory.

Keywords
Augmented Reality, Ubiquitous Computing, Smart Environments, Industrial Design, Kitchen.

IUI'05, January 9–12, 2005, San Diego, California, USA.
ACM 1-58113-894-6/05/0001.

1. INTRODUCTION
Task-oriented spaces benefit from augmented reality projection to inform and direct tasks, especially between multiple tasks in the space. The kitchen is a typical laboratory space where multiple, complex tools are used in time-sensitive routines. Overlaying the space with digital interfaces can help to make cooking more easy, safe and efficient. *CounterActive* teaches cooking through a fixed multimedia projection of step-by-step recipes [3]. The *Anywhere Projector* augments reality by projecting interfaces on any surface in a single room [4]. We have shown a variety of interfaces projected onto the objects and surfaces of the kitchen can aid in the task of cooking, communicating and orienting tasks [2]. We built an augmented reality kitchen where simple sensing and projection augment the refrigerator, cook-top, cabinets, sink and drawers (see Figure 1) [1]. User evaluations of this environment reveal that the *CounterActive* and *Anywhere Projector* models have limited interface modalities and are inflexible in their approach to tasks. A single recipe can be performed in a number of different ways, especially by users of varied experience. Information projection benefits every step if the interface modalities are carefully tailored to the location, expertise and performance of the user(s).

Figure 1. Information projection on the refrigerator (1), the range (2), the cabinet (3), the faucet(4) and drawers(5).

To design an augmented reality interface useful both for novices as well as experts, we developed a structure for modeling tasks in multiple dimensions to account for various degrees of expertise and multiple correct means of execution. Every cooking task is mapped according to the various ways it can be performed according to its possible location, timing and expertise or quality. These interfaces can be especially beneficial when multiple tasks need to be performed in a time-critical manner, such as when someone is preparing a meal. The augmented reality interface can thread the tasks according to when and where certain steps must occur in order that everything be ready at the same time.

2. IMPLEMENTATION

The Augmented Reality Kitchen consists of a typical residential kitchen instrumented with vision-based sensing that projects multi-modal interfaces on the surfaces of the space. We conducted two user evaluations of this kitchen to determine the proper modalities and design for user interfaces to provide useful information for typical kitchen users. In both studies, participants were asked to familiarize themselves with the contents of cabinets and the functioning of the cook-top to reduce the impact of their unfamiliarity with the space, a condition not normally present in typical kitchens. In both cases, CCD cameras above the cabinets and countertops allowed the projected interfaces to act as touch screens while helping to keep track of users' locations.

Figure 2. A step-by-step recipe with camera-based touch interface.

2.1 User Studies

In the first study, a *Counteractive*-style step-by-step recipe directed users through the simple process of soft-boiling an egg (see Figure 2). Self-illuminating drawer handles that indicated the proper location of dishes and ingredients were found to significantly improve performance, despite the fact that users had already inspected the contents of those drawers. The control group had the advantage of being able to consult a paper recipe and at once read and prepare for multiple steps. Some control users were able to, for example, find the tools for multiple steps in the same cabinet, while experimental users patiently waited for each step to be complete before moving on to the next.

For the second study, users performed a series of simple but time-sensitive tasks to prepare a snack tray. Although the augmented reality system helped to prioritize tasks in order to make efficient use of time, the study revealed limitations in the modality of interface. Some control users performed just as well with a paper recipe as experimental users, whereas some experimental users were not able to complete individual tasks. The microwave popcorn called for in this study revealed wide gaps in knowledge between users who had made it before and those who hadn't. This study reveals the vastly different levels of expertise and problem-solving ability in kitchen users leading to different useful modalities of interface.

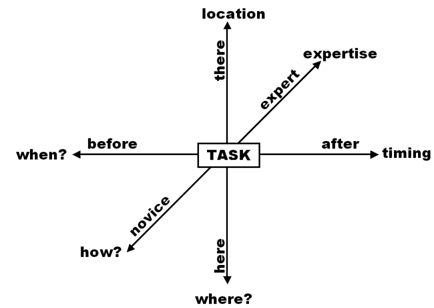

Figure 3. Task mapping in three dimensions: timing, location, and expertise.

2.2 Task Modeling

In order to account for the multiple types of cooking performance even within a single recipe, we propose a models of tasks and users that tailor interface modalities to provide useful information to many types of users. Visual cues based on a linearly depicted task fail in all but the most constrained scenarios. In real life, tasks can be achieved in a number of different ways according to multiple criteria. Many of the steps of a task must be considered together for users to make intelligent decisions about the overall process. Furthermore, the location where each task is performed should be understood at once for users to avoid inefficient repetitive movements. Finally, each task can be performed to varying degrees of quality depending on factors like expertise and available resources. In order to assist such a complex scenario, we map the tasks in three dimensions: time, space, and quality (see Figure 3).

Figure 4. Time map of microwave popcorn preparation indicating periods of activity and pauses.

We begin by mapping typical cooking-related tasks to time and space. First, we map each task according to a flexible timeline that explains the level of activity and thus concentration required at various points in the progress (see Figure 4).

Second, the task is mapped according to the location of necessary tools, ingredients and work surfaces. These can be numerous since often, several utensils can be used to perform the same task. Finally, we apply these task mapping to a multi-modal output that considers user expertise and the desired quality of the outcome. At the beginning, the system provides only expert information: minimally invasive colored projection to indicate temperatures and item locations (see Figure 5).

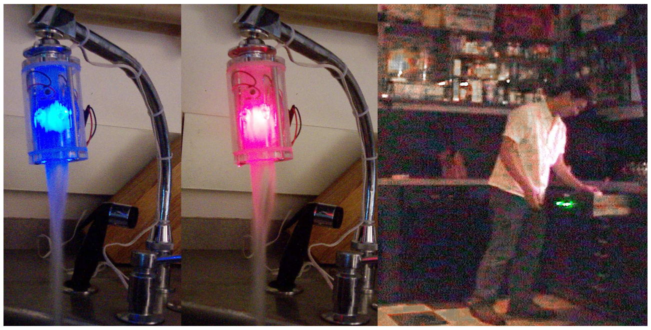

Figure 5. Colored illumination as a non-invasive display for temperature (left) and location (right).

Then, when sensors indicate that users are not progressing, increasingly informative interface modalities are projected until the task progresses smoothly. First, graphics seek to orient users who may be looking in the wrong place. Then, textual information complements the graphics (see Figure 6). In some cases, images and videos are used to show examples of correct procedures. Finally, audio feedback is provided – especially in potentially dangerous situations – to direct a user's attention in a particular direction. Users requiring additional assistance can access "help" files directing basic tasks such as operation of appliances or chopping technique. Only in the case of textual or informational projection is the camera-based touch interface employed.

Figure 6. Three modalities for location-sensitive display: image, text and colored illumination.

2.3 Task Threading

Experienced users can benefit from an augmented reality interface that helps to "thread" multiple tasks by overlapping them in time and space. The temporal and spatial mapping of each task is combined to determine the most efficient use of space and time to prepare a series of items at the same time, such as when cooking a meal (see Figures 7 & 8). When two tasks can be combined into one, the augmented reality interface projects a textual reminder. This can vastly reduce the repetition in a process and keep users from, for example, opening the refrigerator as often.

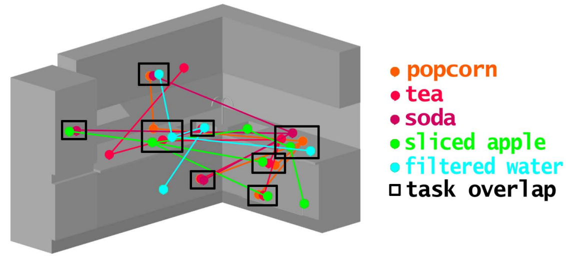

Figure 7. Spatial mapping of multiple tasks reveals overlaps.

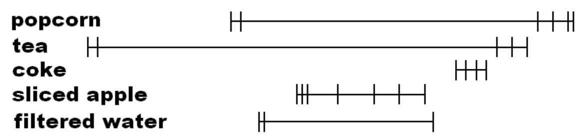

Figure 8. Temporal mapping of multiple tasks allows multi-tasking at times of low intensity.

3. CONCLUSION

Augmented reality projection can greatly enhance the functionality of task-oriented spaces such as the kitchen. We seek to design the interface that informs and choreographs the tasks of preparing food to make the process more easy, safe and efficient. Whereas linear representations of tasks have a limited utility, a system that considers multiple parameters of tasks and represents them through multi-modal interfaces can serve the needs of both novice and expert users. By modeling tasks according to their temporal and spatial characteristics, and modulating the interfaces based on the performance of users, we can provide non-intrusive information to inform and organize tasks. The existing system will be implemented to allow any recipe to be performed with the augmented reality interface able to modulate information based on performance parameters. By structuring the interface in a generalized way, we hope to apply these principles to diverse task-oriented spaces to make tasks more easy and efficient.

4. ACKNOWLEDGMENTS

Our thanks to the MIT Media Lab Counter Intelligence Group and our sponsor Cleanup Corp.

5. REFERENCES

[1] Bonanni, L., Lee, C.H., Gens, R., Selker, T. "Augmented Reality Kitchen: Task-Specific Projection in a Multi-User Work Environment." *in UIST '04 CD-ROM.*

[2] Bonanni, L. and Lee, C.H. "The Kitchen as a Graphical User Interface." *in SIGGRAPH 2004 Electronic Art and Animation Catalog*, 109-111.

[3] Ju, W. et. al. "Counteractive: An Interactive Cookbook for the Kitchen Counter," *in Extended Abstracts CHI 2001*, 269-270.

[4] Pinhanez, C. "Augmented Reality with Projected Interactive Displays." *in Proc. VAA 2001.*