“Let There Be Light”
Examining Interfaces for Homes of the Future

Barry Brumitt & JJ Cadiz

Microsoft Research, One Microsoft Way, Redmond, WA, USA

{barry; jjcadiz}@microsoft.com

Abstract: As smart devices proliferate in the home and become universally networked, users will experience an increase in both the number of the devices they can control and the available interaction modalities for controlling them. Researchers have explored the possibilities of speech- and gesture-based interfaces in desktop scenarios, but it is also necessary to explore these technologies in the context of an intelligent environment, one in which the user’s focus moves off of the desktop and into a physically disparate set of networked smart devices. We report results from a lab study where six families tried various interfaces for controlling lights in the home. Results indicate a preference for speech-based interfaces and illuminate the importance of location awareness.

Keywords: Home automation, multimodal interfaces, speech interfaces, intelligent environments

1 Introduction

There is an ongoing trend towards increasing amounts of technology in the home. The PC, in particular, has moved from something only owned by the enthusiast to a device found in over 50% of US homes. Beyond PCs, “smart” devices such as security cameras, remotely controllable lighting, and mobile phones are also found with increasing prevalence in the home. With the advent of technologies such as Bluetooth, as well as proprietary RF and powerline standards, these smart devices are becoming increasingly connected and, as this trend continues, will be able to behave in a coordinated fashion. We see the first signs of this trend in specialized devices, such as an MP3 player that can broadcast to the stereo in another room over phone lines.

While current devices tend to perform only standalone functions (like the MP3 player or security cameras mentioned above), the increasing connectivity of devices should allow more complex interactions. For example, a camera used for infant monitoring might be able to direct its output to whatever display is most conveniently located for the parent, or be used for another task altogether, such as videoconferencing. This is very different from the current paradigm where each input device can connect to a small fixed set of devices in the home. An environment that has this kind of awareness of its users and an ability to maintain a consistent coherent interaction with the user across a number of heterogeneous smart devices is called an Intelligent Environment [Cohen, 1998], a type of Ubiquitous Computing [Weiser, 1993] system.

Since intelligent environments typically contain many diverse input devices, off-the-desktop multimodal interfaces are a natural result. In this domain, multimodal interfaces refer not only to using more than one modality simultaneously for a given task, but also to enabling the selection of any of multiple modalities for the task. Intelligent environment research systems [Brumitt et al, 2000; Cohen, 1998; Wren et al, 1999] have used diverse interface modalities (including speech, gesture, GUIs, remote controls, etc.), but research in this area has tended to be focused on what is technologically feasible instead of on which modalities users prefer. Intelligent environment researchers haven’t explored what makes a usable home interface.

Unfortunately, desktop PC usability is not necessarily the same as intelligent environment usability. Interface designers cannot assume that all users will have the same interaction affordances, i.e. a keyboard, mouse, and screen. In one room, a touch screen might be available, but only a speech interface in another. Furthermore, while the PC and its peripherals are implicitly co-located (on the same desk, for example), in a home environment the devices are located throughout the home.

One of the barriers to finding out what makes a usable home system is that conducting valid lab and field studies is difficult. Most usability labs weren't built to create a home environment, and field studies with fragile technology in the home can be complex. Furthermore, creating technology robust enough to test can be difficult given the integrated systems built by intelligent environment researchers.

Thus, this research addresses the issue of usability in the home via a low cost "Wizard of Oz" lab study of how users might control lights in a home of the future. Given the substantial effort required to build intelligent environment systems, our hope is that these user data guide researchers toward developing systems that have the most potential to be usable and desirable. In particular, indications of which interaction modalities are most important to usability can drive the selection and development of costly and complex perception systems, such as those that understand speech, interpret physical gestures or determine people's or objects' locations.

In the following section, we discuss previous research on the home environment. In section 3, we introduce our focus on lights and describe different light control interfaces. Section 4 discusses the methodology of our study while section 5 presents our results. We discuss how our findings may direct home automation research in section 6.

2 Related Work

Some previous research exists on interfaces in the home, but its focus has been on specific tasks for home PCs instead of integrated intelligent environments [Plaisant and Shneiderman, 1991; Tetzlaff et al, 1995]. More applicable to this research are several ethnographic field studies of technology use in the home. These studies have focused on the use of television set-top-boxes in the home [O'Brien et al, 1999], the use of tablet PCs in the home environment [McClard and Somers, 2000], and the home environment in general [Mateas et al, 1996]. Venkatesh [1996] complements this work with a theoretical framework of household-technology interaction, based on research beginning in the 1980's of technology adoption in the home.

These studies outline several concepts to keep in mind when designing technology for the home. Most important is the idea that "the implicit design assumptions of the personal computer are inappropriate for the home" [Mateas et al, 1996]. The environments and tasks in the home are different from what is found in the workplace, as are the relationships. For example, Venkatesh [1996] outlines several environments in the home where a variety of tasks take place, ranging from food preparation to watching movies.

The different environment of the home also brings up the issue of form factor. Studies of tablet PC use in the home found that the top reason users enjoyed the tablet PC was that it was portable [McClard and Somers 2000]. Ethnography research reported in [Mateas et al, 1996] also supports the idea of using smaller, portable devices in the home. They write, "The data imply that ubiquitous computing, in the form of small, integrated computational appliances supporting multiple collocated users throughout the home, is a more appropriate domestic technology than the monolithic PC." These studies suggest that people may find integrated intelligent environments to be quite useful for their homes.

In fact, one study that focused on people with special needs found that hand gestures were an acceptable interface, particularly for light control [Machate et al 1999]. Additionally, they found that users adapted quickly to nontraditional interaction modalities.

But overall, little research exists to tell us exactly what kind of interfaces would be best suited for integrated, intelligent, home environments. This paper complements previous work by presenting a lab study of different potential interfaces for intelligent environments in the home.

3 Focusing on Lights

Lights are devices that are universally well understood and have been a typical target of today's nascent home automation systems. Their simplicity and ubiquity drove the selection of light control as a focus for this study. Though many interaction modalities have been proposed for intelligent environments, there has been no explicit comparison of these alternatives for a single task. This paper reports on a lab study of light control usability in an intelligent environment.

Currently, lights are typically turned on via a switch located on the wall or on the light itself. However, with the addition of speech and vision systems, other options become viable. A person might talk or make gestures to control lights. Small screens with access to the position of the user and all lights in the room can provide dynamic, context-sensitive interfaces. Following are some non-

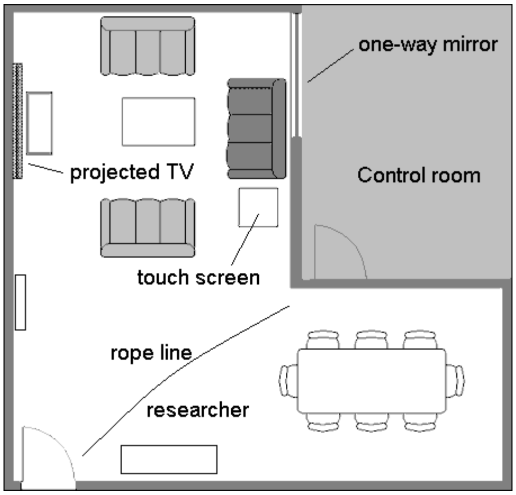

Figure 1: Layout of the lab where the study was conducted.

traditional mechanisms for controlling the lighting in a room, along with descriptions of each:

Plain text computer display: Using any display and point-and-click technology, a list of lights can be displayed. Using this list, users may control lights via slider bars or similar widgets. However, one major issue with any such display is creating consistent, clear labels to describe individual lights.

Graphical computer display: If enough knowledge of the room's physical layout is available, the problem with creating consistent labels for lights can be resolved by showing the user a map of the room with all the lights in it. Furthermore, knowledge of the location of the display and the position of the user can be used to orient the map appropriately.

Voice only: If the room can hear and understand speech, users can control lights by making voice commands. However, this method suffers from the same problem as the plain text computer display: in what way should users refer to individual lights?

Voice with location: One way people can refer to individual lights without using a specialized vocabulary is with commands such as, "Turn on the light to my left." For these commands to work, systems must be imbued with knowledge of where the user is located. A vision system [Krumm et al, 2000] or an active badge system [Ward et al, 1997] could be used to gain this information.

Voice with location and gestures: Vision systems can also be used to recognize gestures [Jojic et al, 2000]. With gesture recognition, users can successfully use multimodal commands to control the lights: for example, saying "Turn on that light," while pointing at a particular light.

Automatic behaviors: Perhaps the best interface is one where the user doesn't have to make any commands at all. Neural net systems may be able to infer the appropriate lighting based on the actions of the user [Mozer, 1998]. For example, if a person enters a room or sits on a particular chair, then the lights should go on to provide appropriate illumination. A Smart Light Switch has also been proposed [Cooperstock et al, 1997] that combines a motion detector with a switch to allow manual overrides of the automatic actions.

Lighting control in the home does not currently present itself as an active problem domain: most people have little difficulty controlling the lights in their home. However, the trend towards increasing numbers of remotely controllable devices in the home indicates a need to examine the modalities users prefer to use when controlling such devices. As soon as there are multiple lights (or other devices) in the room, the problem arises of mapping the control of those devices to some set of input. By examining users preferences for controlling devices that are well understood, we can hope to generalize to other devices in the home. For this reason, we have focused on modalities that are likely to be generalizable (GUI, speech, etc), and avoided those that are not (The Clapper, specialized switches, etc.)

4 Method

To test these methods of controlling lights in the home, we recruited six families, which are typical "users" of home environments. Families consisted of one parent and two children or two parents and one child, and no family members had any experience with a home automation system. A total of 10 parents and 8 kids participated in this study, and all family members received a free software product for their time. Parents ranged in age from 39 to 57 while children ranged in age from 10 to 18. Although families were recruited for this study, participants did the tasks for the experiment individually.

4.1 Physical Setup

For the study we used the Microsoft Research EasyLiving lab. (The EasyLiving project at Microsoft Research [Brumitt et al, 2000] is concerned with architectures for intelligent environments and has built various methods for users to interact with devices in the space.) The lab is a mock-up of a small home living room (Figure 1). One portion of the lab (the lower right, in Figure 1) is set up as a conference room, which was roped off and not used for this study. The remaining part of the lab consists of an entertainment center, three couches, a center table, and some shelves. A television screen is projected on the front wall from behind the one-way mirror.

Participants manipulated the lights in the room using each of the six methods discussed in the previous section, with the exception of "automatic behavior", as it's difficult to "control" lights with automatic interfaces. Users controlled 14 lights for this study. All lights were of the small spotlight type and were mounted in the ceiling in locations shown in Figure 2 (bottom). The two computer display interfaces were made available (one at a time) via a touch screen that was placed on a table in the room.

The methods of light control that required speech and/or gesture recognition were conducted using the "Wizard of Oz" technique: although participants were told that the computer was responding to their commands, a researcher sitting behind the one-way mirror controlled the lights.

To do the tasks for the study, participants first put on a wireless microphone so the "Wizard" could hear them (although we didn't tell them that this was the reason). Participants started by the door to the lab (the lower left portion of Figure 1), and a researcher sat behind the rope line to give the participant instructions.

4.2 Study Scenario

Participants were told that they had agreed to help a friend set up her basement for a surprise birthday party. Participants were also told that the friend recently moved into a new intelligent home that had automated systems for controlling several things, including the lights.

To set up the basement, participants were told that they needed to do two tasks: first, they needed to turn on all the lights in the basement; second, they needed to adjust the lights to create a spotlight effect on the center coffee table, which is where the birthday cake and presents would be placed. The spotlight effect was to be achieved by dimming the lights around the edge of the table while leaving the light above the coffee table at full brightness. Participants were told that they would do these two tasks several times, each time using a different method of controlling the lights.

4.3 Conditions

The first time participants did the tasks, they were told nothing about the room except that it was "intelligent." This task was used to see what people expected of a futuristic intelligent home. When participants started this task, the touch screen was placed in screen saver mode and displayed nothing

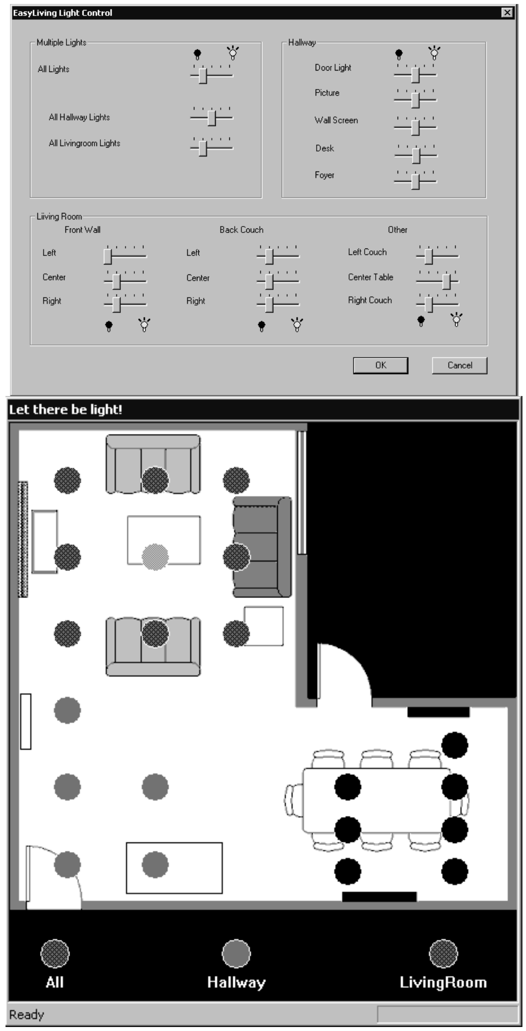

Figure 2: The two touch screen interfaces used for this study. On the top is the text interface, on the bottom is the graphical interface. In the bottom display, tapping lights would cycle them through 0%, 25%, 50%, 75%, and 100% brightness. Participants could not control the lights in the lower-right portion of the room (the conference room).

so that it would not be an obvious avenue to complete the task.

Second, participants did the tasks for all of the methods listed in Section 3 (except the method they used for the very first condition). Ordering of the conditions was balanced such that half of the participants used the touch screen interfaces first while the other half used the speech interfaces first. For the touch screen interfaces, participants were shown the screen with one of the interfaces and asked to do the tasks. The two touch screen interfaces are shown in Figure 2. For the speech interfaces, participants were told to do the tasks with the knowledge that the computer could hear them. If participants did not spontaneously gesture or refer to their location, participants were next shown the three video cameras and told that they were "computer eyeballs." Participants were then asked to do the tasks with the knowledge that the computer could see and hear. If participants did not gesture spontaneously, they were told that the computer could understand "gestures" (researchers were careful not to say "pointing") and asked to try the tasks again. After using each method of controlling the lights, participants indicated on a 5-point Likert scale how much they liked the method and how easy they thought the method was to use.

For the last condition, participants were told they could use any of the methods to control the lights. This condition was used to probe participants' preferences after they knew everything the room could do.

After completing all of the above, participants were shown how automatic lighting features could work in homes. They were asked to sit down on a couch. When they did, the lights above them turned on. When they got up and left the area, the lights turned off.

5 Results

For the first condition when participants were told nothing about the room except that it was "intelligent", we were convinced that a substantial number of people would immediately try to talk to the room to control the lights. However, only one person used speech. Of the remaining sixteen participants, twelve discovered and used the touch screen and four were stopped after trying unsuccessfully for three minutes. With the exception of the one participant who talked to the room, all participants wandered around the room looking for a device that looked like it might control the lights.

5.1 Ratings of Liking and Ease of Use

After using each method of controlling the lights, participants rated on a 5-point Likert scale how much they liked the interface and how easy they thought the interface was to use. The results of these data are shown in Table 1.

Mann-Whitney U tests were used to test for differences in these data. Overall, the only significant differences were found when comparing some of the speech interfaces to the touch panel interfaces. On the measure of liking, the speech interface was rated significantly higher than the touch screen graphical interface ($z = 1.97$, $p = .049$) and the text interface ($z = 2.7$, $p = .006$). The speech with gesture interface was also rated significantly higher than the touch screen text interface ($z = 2.8$, $p = .006$). On the measure of ease of use, the touch screen text interface was rated significantly lower than both the speech interface ($z = 2.1$, $p = .04$) and the speech with gesture interface ($z = 2.8$, $p = .005$). These data indicate that speech was the preferred interface.

5.2 Choosing a Method for the Final Condition

For the last condition when participants could use any method of controlling the lights they desired, fifteen people used speech and only two people used the touch screen, indicating that speech was the preferred interface. Of the fifteen people who used speech, eight used only their voice while seven used a combination of their voice and some type of gesture. Additionally, nine participants used speech vocabulary that assumed the system possessed knowledge about the location of things in the room, e.g. "Turn on the lights around the center table" or "Dim the ones behind the couch over here." Both people who used the touch screen used the graphical interface.

5.3 Vocabulary for Controlling Lights

When trying to control lights, people's speech commands have two elements: an action (on, off, etc.), and a reference to a particular light or group of lights. The vocabulary used for the action part of the command is quite small. When analyzing transcripts from the study, we found that just four words—"on", "off", "bright", and "dim"—were used in 81% of the 198 commands.

	I liked this interface 1 = hated it 5 = loved it	**This interface was easy to use 1 = very difficult, 5 = very easy**
Touch screen – graphical interface	4.0	4.0
Touch screen – text interface	3.0	3.5
Speech	5.0	5.0
Speech with location	4.0	4.0
Speech with gesture	5.0	5.0

Table 1: Median scores (higher is better) of people's reactions to each method of controlling the lights.

Unfortunately, the vocabulary used to refer to lights wasn't nearly as predictable. Of the 198 speech commands, 16% used a reference to an object to refer to a light ("the light above the couch"), 30% used an area reference to refer to a light ("all the lights in the living room"), 23% included relative terms such as "left", "right", "back", and "front", and 18% used an indirect reference ("this light", "all on this side", "over there in the corner"). Notably, all references to relative terms were given relative to the television, e.g. "the left couch" referred to the couch to the left of a person facing the television screen.

Thus, while interpretation of the action part of the command is straightforward, interpretation of the object reference is more complex.

5.4 Looking at Lights

There are numerous ways to address the issue of resolving which object a person is referencing. One can label all the lights, but this requires a person to learn a naming scheme for all the lights in the house, and visitors will be faced with the problem of trying to guess the labels selected by the owner. One can use vision so that people can point to lights, but pointing was not completely intuitive to people (only 2 of the 15 people whose data we could analyze pointed at lights before we told them that the computer could recognize gestures).

However, as we observed participants during the study, there was one action that everyone seemed to do when referring to lights: they looked at the light they wanted to control. In fact, in only 9% of the tasks did people never look at the light they were controlling for any of the commands they issued. In 25% of the tasks, people looked at the light during some of their commands, and in the remaining 66%, people always looked at the light they wanted to control.

Gaze tracking has been previously proposed as a reinforcing modality to speech [Sarukkai et al, 1997]. This study demonstrates that for scenarios of device control in the home, gaze may be fruitfully used to prune the list of devices to which the user may be referring.

6 Discussion

Before discussing the results of this study at length, we should note its shortcomings. Mainly, this research has the generalizability issues in common with other lab studies. We were measuring participants' reactions to different methods of home automation based on interacting with the room for only a few minutes. Opinions about the different methods of controlling lights might be more pronounced if we were to install each of the methods in the participants' homes for a month and then ask for their reactions. Participants were giving us their initial reactions to the technology, thus a novelty effect may have influenced the data. However, first reaction data to new technology is valuable since participants began the study with very few, if any, preconceived notions about how home automation systems might work.

Investigating speech as an input modality presented several difficulties. We did not control for social effects; for example, people might not prefer speech if other human listeners were present. We also did not limit the acceptable vocabulary, or otherwise emulate the current state-of-the-art in speech systems. The goal was to allow the comparison between modalities independent from any technological constraints.

6.1 Which was the Preferred Interface?

The data indicate that participants prefer to use their voice to control lights. Speech interfaces had the highest initial reaction scores, and when participants were given the choice of using any method of controlling the lights, they chose to use their voice.

While this finding is notable, it is possible that this result is due to the novelty effect. Additionally, all tests were done with only one person in the room. In a social setting, speech may not receive the same high marks. The more interesting findings are the choice of vocabulary and the coupling of speech with gaze and location detection.

6.2 Does Adding Gesture or Location Recognition Help?

Given the research concentrating on systems that can see a person's location or recognize gestures, it's important to analyze whether, from a user's point of view, adding these systems helps. The data indicate that users don't perceive any added value in a system that is able to see them. Ratings of speech interfaces vs. speech interfaces with location awareness vs. speech interfaces with gesture recognition were essentially the same. In fact, during the study, people didn't seem to perceive much of a difference between the conditions, sometimes doing the task in the same way (just using voice commands) even though we told them that the computer could see them and understand gestures.

However, we don't believe this means that homes of the future only need to be outfitted with microphones, or that vision systems research for home automation is misguided. People liked speech because it worked nearly perfectly, and it worked nearly perfectly because we had a "wizard" (another researcher) controlling the lights to make sure that they responded as perfectly as possible to user requests. But after observing how people tried to talk to lights, we believe a system would not be able to function well without the use of location-aware technology. This is primarily due to the difficulty in interpreting the vocabulary used to describe particular devices. As noted in Section 5.3, references depend on knowing the location of the devices and of other objects in the room.

A speech-driven device control system could function better still if gaze tracking were available, as the modalities can reinforce each other. From our observations, the best "gesture" to recognize is the place in a room where someone is looking when they utter a command. This approach has been proposed by [Kuno et al, 1999] for desktop computers, and [Oviatt 1999; Oviatt 2000] has studied the use of multimodal interfaces to reduce speech recognition error rates with pen-based computers. Oviatt [2000] reported an improvement of as much as 41% when using multimodal architectures to recognize speech.

Thus, for intelligent environments, perception systems may not be best used to add more features (such as gesture recognition). Rather, perception systems that deliver gaze or gesture may be best used in conjunction with speech recognition systems to interpret people's commands as correctly as possible. In any case, location awareness is essential for these systems to function acceptably.

6.3 Evaluating System Designs

Based on this study, the need for some degree of location awareness in an intelligent environment is clear. Users prefer speech- and screen-based UI's, which necessarily utilize location information. Location-aware systems can be built to provide information at varying resolutions, from providing information about the building in which the user is located, down to providing the exact (x,y) location of the user in a room. We believe a fruitful area for system design lies in providing awareness of which room a user is in.

It is possible to sense in which room a user is located by using a beacon-based location detection technology (e.g. [Want et al, 1993]). Since the locations of lights are typically fixed, it is not unreasonable to expect this information to be hand-entered into a world model. However, the location of the user is dynamic, and therefore must be actively perceived. By knowing the location of the input device (microphone or screen), a fixed model can provide a list of likely devices that are to be controlled by simply determining the devices that are proximate to the input device, and therefore, also to the user. Study participants rarely employed location references given relative to themselves, and so perceiving their exact location is of only slight usefulness. Of course, if mobile input devices are considered, their location must be sensed in order to provide these same benefits as for fixed devices. A beacon-based system would provide most of these benefits.

The alternative to a beacon-based system is a triangulation-based one (e.g. via ultrasonic sensors [Ward et al, 1997] or cameras [Krumm et al, 2000]), which provide more precise geometric localization. In fact, cameras alone can provide location estimation, gesture understanding, and gaze detection. However, it is necessary to recognize the substantial cost and research effort associated with these higher-fidelity options along with a diminishing rate of returns in terms of increased performance.

Overall, though gaze and gesture have been suggested for these scenarios, their benefits relative to costs seem small. We believe this may also be the case for triangulation-based localization methods. Though they would allow more complex behaviors (such as lights coming on when the person sits on the couch), these seem to be less important features.

7 Concluding Remarks

We have reported findings from an initial study of interfaces for the home. Based on our observations of families using a variety of interfaces in the home, we observed several trends. First was a preference for speech-based interfaces, even though participants did not expect this modality to be available. Second, we observed that for speech- or screen-based interfaces to be effective, the system must posses location awareness of at least room-level fidelity. Third, gaze is a useful modality to aid in disambiguation with speech interfaces. Finally, gesture recognition is perceived as being of only moderate usefulness to users in this environment.

Future work should examine the use of speech for other device control tasks in the home to ensure that the results presented herein generalize and are sustained over the long term. Additionally, study of the vocabulary used to describe devices and actions would be helpful. Finally, the hypothesis that room-level location awareness is sufficient should also be tested.

Acknowledgements

Thanks go to Kris Keeker, Steven Shafer, Mary Czerwinski, Brian Meyers, Michael Hale, and Amanda Kern for their assistance in study design, UI implementation, test coordination, and data analyses.

References

Brumitt, B., Meyers, B., Krumm, J., Kern, A., and Shafer, S. (2000). EasyLiving: Technologies for Intelligent Environments. Proc. of the Intl. Conf. on Handheld and Ubiquitous Computing 2000, 12-27.

Cohen, M. (1998). Design principles for intelligent environments Proceedings of the AAAI Symp. on Intelligent Environments, 36-43.

Cooperstock, J.R. Fels, S. S., Buxton, W. and Smith, K.C. Reactive Environments: Throwing Away Your Keyboard and Mouse. Comm of the ACM, Vol. 40, No. 9, September 1997.

Jojic, N., Brumitt, B., Meyers, B. Harris, S., Huang, T. (2000). Detection and Estimation of Pointing Gestures in Dense Disparity Maps. Proc. of the 4th Intl. Conf. on Automatic Face and Gesture Recognition, 468-475.

Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M., and Shafer, S. (2000). Multi-Camera Multi-Person Tracking for EasyLiving. IEEE Workshop on Visual Surveillance, 3-10.

Kuno, Y., Ishiyama, T., Nakanishi, S., and Shirai, Y. (1999). Combining Observations of Intentional and Unintentional Behaviors for Human-Computer Interaction. Proc. CHI ‘99, 238-245.

Machate, J., Bekiaris, E., Nikolaou, S. Control it by Gestures: Results from a Wizard of Oz Experiment in a Smart Home Environment. Proc. of the 5th European Conf. for the Advancement of Assistive Technology (AAATE 99), Düsseldorf, Germany, November 1999.

Mateas, M., Salvador, T., Scholtz, J., and Sorensen, D. (1996) Engineering Ethnography in the Home. Conference Companion for CHI ’96, 283-284.

McClard, A., and Somers, P. (2000). Unleashed: Web Tablet Integration into the Home. Proc. CHI 2000, 1-8.

Mozer, M. (1998). The Neural Network House: An Environment that Adapts to its Inhabitants. Proc. AAAI Symp. on Intelligent Environments, 110-114.

O’Brien, J., Rodden, T., Rouncefield, M., and Hughes, J. (1999). At Home with the Technology: An Ethnographic Study of a Set-Top-Box Trial. ACM Trans. on CHI, 6(3), 282-308.

Oviatt, S. (1999). Mutual Disambiguation of Recognition Errors in a Multimodal Architecture. Proc CHI ’99, 576-583.

Oviatt, S. (2000). Taming Recognition Errors with a Multimodal Interface. Comm. ACM. 43(9), 45-51.

Plaisant, C., and Shneiderman, B. (1991). Scheduling On-Off Home Control Devices. Conference Companion for CHI ’91, 459-460.

Sarukkai, R.R. and Hunter, C. (1997) Integration of Eye Fixation Information with Speech Recognition Systems. In Proc. of the European Conf. on Speech Communication and Technology, 97-100.

Tetzlaff, L., Kim, M., and Schloss, R. (1995). Home Health Care Support. Conference Companion for CHI ’95, 11-12.

Venkatesh, A. (1996). Computers and Other Interactive Technologies for the Home. Comm. of the ACM. 39(12), 47-54.

Want, R, Hopper, A. (1992) Active Badges and Person Interactive Computing Objects, IEEE Trans. On Consumer Electronics, Vol, 38, No. 1, 10-20.

Ward, A., Jones, A., Hopper, A. (1997). A New Location Technique for the Active Office. IEEE Personal Communications, 4(5) ,42-47.

Weiser, M. (1993). Some Computer Science Issues in Ubiquitous Computing. Comm. ACM. 36(7), 75-84.

Wren, C., Basu, S., Sparacino, F., and Pentland, A. (1999). Combining Audio and Video in Perceptive Spaces. Proceedings of the 1st International Workshop on Managing Interactions in Smart Environments, 44-55.