Experiments In Interaction Between Wearable and Environmental Infrastructure Using the Gesture Pendant

Daniel Ashbrook, Jake Auxier, Maribeth Gandy, and Thad Starner
College of Computing and Interactive Media Technology Center
Georgia Institute of Technology
Atlanta, GA 30332-0280 USA
{anjiro, jauxier1, maribeth, thad}@cc.gatech.edu

Abstract

The Gesture Pendant, a computer vision system worn as a piece of jewelry, allows the wearer to control electronic devices in the environment through simple hand gestures. Gestures provide an advantage over traditional device interfaces (such as remote controls) in that they are easily used by all people, including those with such disabilities as loss of vision, motor skills, and mobility. The Gesture Pendant can also be used for medical monitoring—as the user makes gestures, the pendant can analyze the movement of the hands to detect certain tremors, the frequency of which can indicate the progress of some diseases such as Parkinson's.

1 Introduction

The goal of the Gesture Pendant is to allow the wearer to control elements in the home via hand gestures. Devices such as home entertainment equipment and the room lighting can be controlled with simple movements of the hand. Building on previous work that used a wearable camera and computer to recognize American Sign Language (ASL) [Starner et al., 1998], we have created a system that consists of a small camera worn as a part of a necklace or pin. The video from the camera is analyzed and pre–defined gestures are detected, which trigger devices connected to various home automation systems. Thus, the wearer can simply raise or lower a flattened hand to control the light level and can control the volume of the stereo by raising or lowering a pointed finger.

2 Motivation and Interaction

Why do we want to use hand gestures to control home automation systems? Automation offers many benefits to the user, especially the elderly or disabled; however, the interfaces to these systems are generally poor. The most common interface to a home automation system such as X10 is a remote control with small, difficult to push buttons with cryptic text labels that are hard to read. This interface also relies on the person having the remote control with them at all times. Portable touchscreens are emerging as a popular interface, however they present many of the same problems as remotes, with the additional difficulty that the interface is now dynamic and harder to learn. Other interfaces include wall panels, which require the user to go to the panel location to use the system, and phone interfaces, which still require changing location and pressing small buttons.

While speech recognition has long been viewed as the ultimate interface for home automation, there are many problems in this domain. In a house with more than one person, a speech interface could result in a disturbing amount of noise, as all the residents would be constantly talking to the house. Also, if the resident is listening to music or watching a movie, she would have to speak very loudly to avoid being drowned out by the stereo or television. Ambient noise can also cause errors in the speech recognition systems.

One potential problem with using gestures to control devices is the need to contrive enough gestures, or having too many gestures to remember. This, of course, defeats the purpose of having a simple interface. To address this problem, we have experimented with making the pendant system context–sensitive. Listed below are several possible combinations of the pendant and context.

1. **The pendant alone**

   With no context sensitivity, each device must have a distinct gesture for every function the user wishes to control. For example, to adjust the stereo volume, the temperature of the thermostat, and the level of lighting in the room, three different gestures would be required.

2. **The pendant and speech recognition**

   By using speech input, the user can select from a set of devices that all share the same gestures. For example, the user could have a single gesture to control the volume, the thermostat and the light. Before performing the gesture, the user would indicate the desired device by speaking its name. This, of course, would present many of the previously discussed drawbacks of speech recognition, except that the system would be less speech dependent than one which was solely speech–controlled.

3. **The pendant and orientation**

   Another way to select between several devices is to use the physical orientation of the pendant. To choose a target to control, the user faces the desired device. This method of selection can be accomplished by placing the transmitter that sends remote control codes on the pendant rather than near the device being controlled. By limiting the direction of transmission, the signals will only reach the device that the wearer is facing. This method, however, can waste battery power, in that it requires the pendant to send remote control codes for all devices associated with that particular gesture, as it has no way of knowing which device the user is actually attempting to control. One solution for this is to place fiducials on each device, and have the pendant camera recognize what is being controlled. Using fiducials also provides an opportunity for device–specific feedback—if the pendant system knows what device it's facing, it can inform the user by illuminating a light on the device, for example. The biggest downside to using orientation as a source of context is that the user must move from device to device; this could obviously be a problem for the elderly or handicapped.

4. **The pendant and location**

   By using room–level tracking, a set of gestures may be defined on a room–by–room basis. This would require the user to move to the room where the target device is located before performing the gesture. For multiple devices in a room, distinct gestures would be required for each. While this scheme still requires the user to move, it may be reasonable to assume that the wearer will not want to control devices in other rooms.

Such sources of context could be combined in any permutation to create an ideal system for the individual. To date we have tested configurations 1, 2 and part of 3 and are exploring adaptation of our software to recognize fiducials.

3 As an enabling technology

All types of people can use the Gesture Pendant. However, the difficult–to–use interfaces of home automation equipment that make the Gesture Pendant useful for healthy adults make the pendant doubly useful for the elderly. Despite any number of physical impairments that may limit the user's motor skills, mobility, or sight, the Gesture Pendant can still be used, especially for assistive tasks such as opening and closing doors, using appliances, and accessing emergency systems.

The same interface problems are faced by those with disabilities such as cerebral palsy and multiple sclerosis. However, a study has shown that even people with extremely impaired motor skills due to cerebral palsy are able to make between 12 and 27 distinct gestures [Roy et al., 1994], which could be used as input to the gesture pendant. Therefore, the Gesture Pendant can be an interface alternative that can allow people who are unable to use some of the more traditional interfaces to take advantage of the independence that home automation can afford them.

4 Medical Monitoring

As a user makes movements in front of the Gesture Pendant, the system can not only look for specific gestures but can also analyze how the user is moving. Therefore, a second use of the Gesture Pendant is as a monitoring system rather than as an input device—as the user makes a gesture, the pendant can detect the presence and frequency of a hand tremor.

As discussed above, the target populations for the Gesture Pendant are the elderly and disabled. Many of the diseases that these populations suffer have a pathological tremor as a symptom. A pathological tremor is an involuntary, rhythmic, and roughly sinusoidal movement [Elble and Koller, 1990]. These tremors can appear in a patient due to disease, aging, and drug side effects; these tremors can also be a warning sign for emergencies such as insulin shock in a diabetic. Currently, we are interested in recognizing essential tremors (4-12 Hz) and Parkinsonian tremors (3-5 Hz) [Elble and Koller, 1990], since determination of the dominant frequency of the tremor can be helpful in early diagnosis and therapy control of such disorders [Hefter et al., 1989].

The medical monitoring of tremors can serve several purposes. The data can simply be logged over days, weeks, or months for use by the doctor as a diagnostic aid. Upon detecting a tremor or a change in the tremor, the user might be reminded to take medication, or the physician or family members could be notified as appropriate. Tremor sufferers who do not respond to pharmacological treatment can have a deep brain stimulator implanted in their thalamus [Hubble et al., 1996]. This stimulator can help reduce or eliminate the tremors, but the patient must control the device manually. The Gesture Pendant data could be used to provide automatic control of the stimulator. Another area in which tremor detection would be helpful is in drug trials. The subjects involved in these studies must be closely watched for side-effects and the pendant could provide day-to-day monitoring.

5 Gesture Pendant Hardware and Software

The motivation behind the Gesture Pendant calls for a small, lightweight wearable device. At first we considered a hat mount, but concluded that gestures would be too hard to recognize if made in front of the body and difficult to perform if made in front of the hat. Due to the off-the-shelf nature of the components (leading to larger size and heavier weight than ideal), we decided that a pendant form was ideal. Using custom-made parts, the hardware could be shrunk considerably, and other form factors such as a brooch or, assuming sufficient miniaturization, a shirt button or clasp could be possible.

Since the goal of the Gesture Pendant was to detect and analyze gestures quickly and reliably, we decided upon an infrared illumination scheme to make color segmentation less computationally expensive. Since black and white CCD cameras pick up infrared well, we used one with a small form factor (1.3" square) and an infrared-pass filter mounted in front of it. To provide the illumination, we used sixteen near-infrared LEDs in a ring around the camera (Figure 1). The first incarnation had a lens with a roughly 90 degree field of view, but that proved to limit the gesture space too much. A wider angle lens of 160 degrees worked much better, despite the fisheye effect. The design of the pendant is similar to the Toshiba "Motion Processor", which uses a camera and IR LEDs as an input to a desktop gesture–based interaction system [Toshiba, 1998].

The eventual goal is to incorporate all components of the gesture pendant into one wearable device; however, for the sake of rapid prototyping we used a desktop computer to do the bulk of the image processing. This also allowed us to centralize the control system, by using standard home automation devices such as the Slink-E (a computer–controlled universal remote) and X10. To send the video to the desktop, we used a 900 MHz video transmitter/receiver pair. The transmitter is powerful enough that cordless 900 MHz phones do not interfere with it, and the receiver can be tuned to a range of channels to avoid conflicting signals from multiple pendants.

Since one of the groups that we feel could most benefit from the Gesture Pendant is the elderly, it is important to make it as unobtrusive as possible. This means it must be inconspicuous, lightweight, and

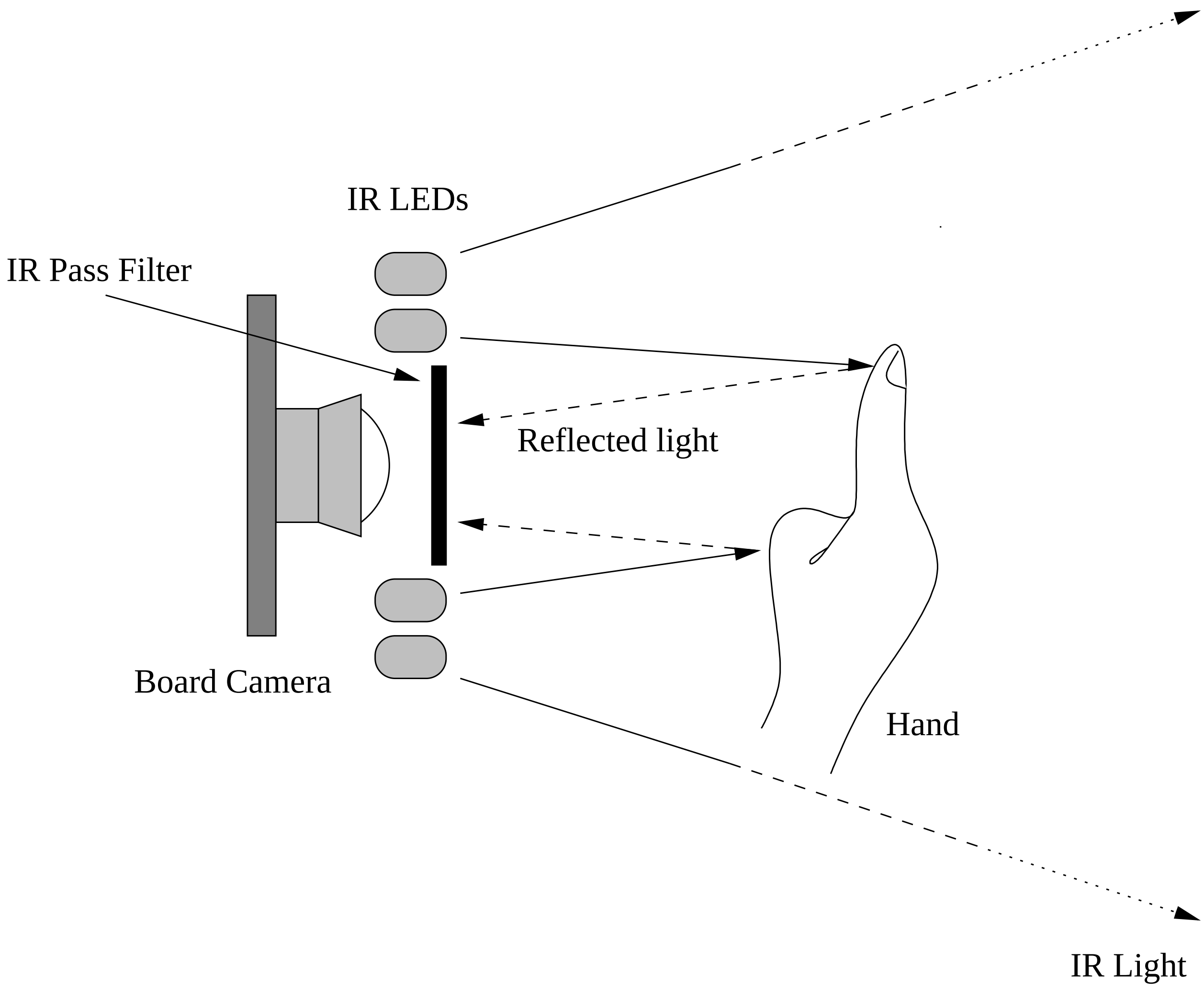

Figure 1: Side view of pendant with IR reflecting off the user's hand

uncomplicated. Also, since the Gesture Pendant is a wearable device, and one that is constantly in full view, it will be important to make it more attractive. The ring of LEDs makes it appear somewhat jewelry-like, but with a smaller form-factor and some principles of design applied to it, it will become more appealing.

The gesture recognition system incorporates two kinds of gestures: control gestures and user-defined gestures. Control gestures provide continuous control of a device that varies with the magnitude of the gesture, and are recognized with an accuracy of 95%. Control gestures would be useful for devices that provide a range of adjustment, such as the volume of a stereo or the brightness of a light. User defined gestures provide discrete output for a single gesture, and are most useful for on–off tasks such as opening a door. User defined gestures are detected with an accuracy of 97%. The gesture algorithm is described in more detail elsewhere [Starner et al., 2000].

6 Future Work

The current implementation of the Gesture Pendant uses a wireless transmitter to send the video data to a desktop PC where it is analyzed and automation commands are issued. The next step in our work is to place all of the computation onto the body in the form of a wearable computer and eliminate the need for a desktop machine.

The monitoring of tremors and motor skills could be expanded to do more complex analyses of the types of tremors in 3D. For example, Parkinson's sufferers often exhibit a complex "pill rolling" tremor, which could be detected and analyzed. We could also determine more characteristics of the user's motor skills, such as slowness of movement or rigidity, that could indicate the onset of stroke or Parkinson's. We could also design the gestures so that, while they would be used to control devices in the house, they would optimally reveal features of the user's manual dexterity and movement patterns.

Another, more advanced use for the pendant in terms of monitoring would be to observe more about the wearer's activities. For example the pendant could take note of when the user eats a meal or takes medication. It could keep a record of the general activity level of the wearer or notice if she falls down. This would further our goal of providing services for the elderly and disabled that allow them increased independence in the home.

7 Conclusion

We have demonstrated a wearable gesture recognition system that can be used in a variety of lighting conditions to control home automation. Through the use of a variety of contextual cues, the Gesture Pendant can disambiguate the devices under its control and limit the number of gestures necessary for control. We have shown how such a device may have enough merit to be used as a convenience by the elderly but also provide additional functionality as a medical diagnostic.

8 Acknowledgments

Funding for the project was provided in part by the Georgia Tech Broadband Institute, the Aware Home Research Initiative, the Georgia Tech Research Corporation, and the Graphics, Visualization, and Usability Center. Special thanks to Rob Melby and David Minnen for the blob tracking software.

References

[Elble and Koller, 1990] Elble, R. J. and Koller, W. C. (1990). *Tremor*. Johns Hopkins UP, Baltimore, MD.

[Hefter et al., 1989] Hefter, H., Homberg, V., and Freund, H. J. (1989). Quantitative analysis of voluntary and involuntary motor phenomena in parkinson's disease. In Przuntek, H. and Riederer, P., editors, *Early Diagnosis and Preventative Therapy in Parkinson's Disease*. Springer-Verlag Wien, New York, NY.

[Hubble et al., 1996] Hubble, J., Busenbark, K., and Wilkinson, S. (1996). Deep brain stimulation for essential tremor. In *Neurology*, volume 46, pages 1150–1153.

[Roy et al., 1994] Roy, D. M., Panayi, M., Erenshteyn, R., Foulds, R., and Fawcus, R. (1994). Gestural human-machine interaction for people with severe speech and motor impairment due to cerebral palsy. In *Conference on Human Factors in Computing Systems*, Boston, MA.

[Starner et al., 2000] Starner, T., Auxier, J., Ashbrook, D., and Gandy, M. (2000). The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In *IEEE Intl. Symp. on Wearable Computers*, Atlanta, GA.

[Starner et al., 1998] Starner, T., Weaver, J., and Pentland, A. (1998). Real-time American Sign Language recognition using desk and wearable computer-based video. *IEEE Trans. Patt. Analy. and Mach. Intell.*, 20(12).

[Toshiba, 1998] Toshiba (1998). Toshiba's motion processor recognizes gestures in real time. Available at: http://www.toshiba.com/news/980715.htm.