WUW - Wear Ur World - A Wearable Gestural Interface

Pranav Mistry
MIT Media Lab
20 Ames Street
Cambridge, MA 02139 USA
pranav@media.mit.edu

Pattie Maes
MIT Media Lab
20 Ames Street
Cambridge, MA 02139 USA
pattie@media.mit.edu

Liyan Chang
MIT Media Lab
20 Ames Street
Cambridge, MA 02139 USA
liyanchang@mit.edu

Copyright is held by the author/owner(s).
CHI 2009, April 4 – 9, 2009, Boston, MA, USA
ACM 978-1-60558-246-7/09/04.

Abstract
Information is traditionally confined to paper or digitally to a screen. In this paper, we introduce WUW, a wearable gestural interface, which attempts to bring information out into the tangible world. By using a tiny projector and a camera mounted on a hat or coupled in a pendant like wearable device, WUW sees what the user sees and visually augments surfaces or physical objects the user is interacting with. WUW projects information onto surfaces, walls, and physical objects around us, and lets the user interact with the projected information through natural hand gestures, arm movements, or interaction with the object itself.

Keywords
Gestural Interaction, Augmented Reality, Wearable Interface, Tangible Computing, Object Augmentation

ACM Classification Keywords
H5.2. Information interfaces and presentation: Input devices and strategies; Interaction styles; Natural language; Graphical user interfaces.

Introduction
The recent advent of novel sensing and display technologies has encouraged the development of a variety of multi-touch and gesture based interactive

systems [4]. Such systems have moved computing onto surfaces such as tables and walls and brought the input and projection of digital interfaces into one-to-one correspondence such that the user may interact directly with information using touch and natural hand gestures. A parallel trend, the miniaturization of mobile computing devices permit "anywhere" access to information, so we are always connected to the digital world.

Unfortunately, most gestural and multi-touch based interactive systems are not mobile and small mobile devices fail to provide the intuitive experience of full-sized gestural systems. Moreover, information still resides on screens or dedicated projection surfaces. There is no link between our interaction with these digital devices and interaction with the physical world around us. While it is impractical to modify all physical objects and surfaces into interactive touch-screens, and to turn everything into network enabled devices, it is possible to augment environment around us with visual information. When digital information is projected onto physical objects, instinctively it makes sense to interact with it in similar patterns as with physical objects: through hand gestures.

In this paper, we present WUW, a computer-vision based wearable and gestural information interface that augments the physical world around us with digital information and proposes natural hand gestures as the mechanism to interact with that information. It is not the aim of this research to present an alternative to multi-touch or gesture recognition technology, but rather to explore the novel free-hand gestural interaction that WUW proposes.

Related Work

Recently, there has been a great variety of multi-touch interaction based tabletop (e.g.[5,6,11,13]) and mobile device (e.g.[2]) products or research prototypes that have made it possible to directly manipulate user interface components using touch and natural hand gestures. Several systems have been developed in the multi-touch computing domain [4], and a large variety of configurations of sensing mechanisms and surfaces have been studied and experimented within this context. The most common of these include using specially designed surfaces with embedded sensors (e.g. using capacitive sensing [5, 16]), cameras mounted behind a custom surface (e.g. [10, 22]), cameras mounted in front of the surface or on the surface periphery (e.g. [1, 7, 9, 21]). Most of these systems depend on the physical touch-based interaction between the user's fingers and physical screen [4] and thus do not recognize and incorporate touch independent freehand gestures.

Oblong's g-speak [12] is a novel touch-independent interactive computing platform that supports a wide variety of freehand gestures to a very precise accuracy. Initial investigations, though sparse due to the availability and prohibitive expense of such systems, nonetheless reveal exciting potential for novel gestural interaction techniques - that conventional multi-touch tabletop and mobile platforms do not provide. Unfortunately, g-speak uses an array of 10 high-precision IR cameras, multiple projectors, and an expensive hardware setup that requires calibration. There are a few research prototypes (e.g. [8, 10]) that can track freehand gestures regardless of touch, but they also share some of the same limitations as most multi-touch based systems: user interaction is required



figure 2. WUW applications

for calibration, the hardware is at reach from the users, and the device has to be as large as the interactive surface, which limits its portability; nor does it project on a variety of surfaces or physical objects. It should also be noted that most of these research prototypes rely on custom hardware and that reproducing them represents a non-trivial effort.

WUW also relates to augmented reality research [3] where digital information is superimposed on the user's view of a scene. However, it differs in several significant ways: First, WUW allows the user to interact with the projected information using hand gestures. Second, the information is projected onto the objects and surfaces themselves, rather than onto glasses or goggles, which results in a very different user experience. Moreover, the user does not need to wear special glasses (and in the pendant version of WUW the user's entire head is unconstrained). Simple computer-vision based freehand-gesture recognition techniques such as [7, 9, 17], wearable computing research projects (e.g. [18, 19, 20]) and object augmentation research projects such as [14, 15] inspires the WUW prototype.

What is WUW?

WUW is a wearable gestural interface. It consists of a camera and a small projector mounted on a hat or coupled in a pendant like mobile wearable device. The camera sees what the user sees and the projector visually augments surfaces or physical objects that the user is interacting with. WUW projects information onto the surfaces, walls, and physical objects around the user, and lets the user interact with the projected information through natural hand gestures, arm movements, or direct manipulation of the object itself.

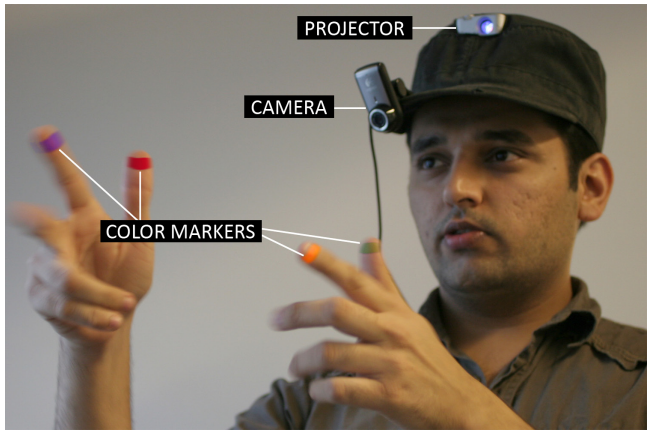


figure 1. WUW prototype system.

The tiny projector is connected to a laptop or mobile device and projects visual information enabling surfaces, walls and physical objects around us to be used as interfaces; while the camera tracks user hand gestures using simple computer-vision based techniques.

The current WUW prototype implements several applications that demonstrate the usefulness, viability and flexibility of the system. The map application (see Figure 2A) lets the user navigate a map displayed on a nearby surface using familiar hand gestures, letting the user zoom in, zoom out or pan using intuitive hand movements. The drawing application (Figure 2B) lets the user draw on any surface by tracking the fingertip movements of the user's index finger. The WUW prototype also implements a gestural camera that takes photos of the scene the user is looking at by detecting the 'framing' gesture (Figure 2C). The user can stop by any surface or wall and flick through the photos he/she has taken. The WUW system also augments physical

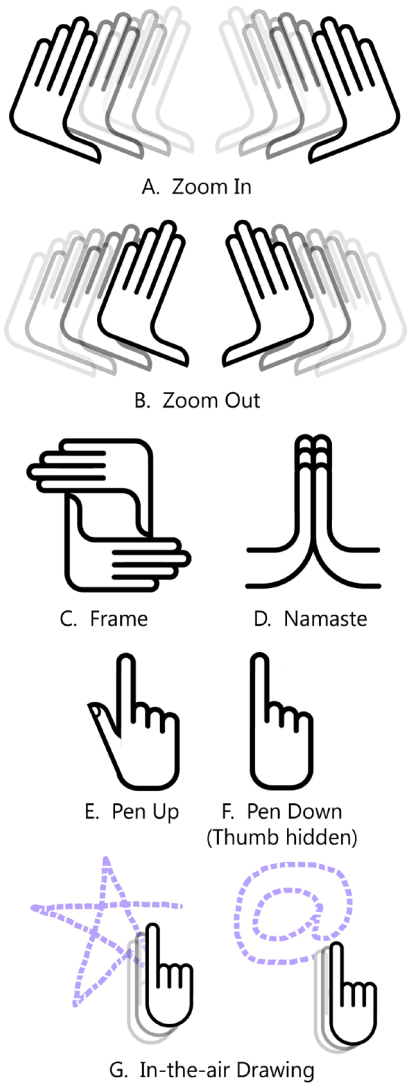


figure 3. Example gestures.

objects the user is interacting with by projecting more information about these objects projected on them. For example, a newspaper can show live video news or dynamic information can be provided on a regular piece of paper (Figure 2D). The gesture of drawing a circle on the user's wrist projects an analog watch (Figure 2E). The WUW system also supports multiple users collaborating and interacting with the projected digital information using their hand movements and gestures.

Gestural Interaction of WUW

The WUW primarily recognizes three types of gestures:

1. Gestures supported by multi-touch systems
2. Freehand gestures
3. Iconic gestures (in-the-air drawings)

Figure 3 shows a few examples of each of these gesture types. WUW recognizes gestures supported and made popular by interactive multi-touch based products such as Microsoft Surface [11] or Apple iPhone [2]. Such gestures include zoom in, zoom out or pan in a map application or flip though documents or images using the movements of user's hand or index finger. The user can zoom in or out by moving his hands/fingers farther or nearer to each other, respectively (Figure 3A and 3B). The user can draw on any surfaces using the movement of the index finger used as a pen (Figure 3E and 3F).

In addition to such gestures, WUW also supports freehand gestures (postures). One example is to touch both the index fingers with the opposing thumbs, forming a rectangle or framing gesture (Figure 3C). This gesture activates the photo taking application of WUW, which lets the user take photo of the scene he/she is looking at, without needing to use/click a camera. Another example of such gestures is the 'Namaste' posture (Figure 3D), that lets the user navigate to the home screen of WUW from within any application.

The third type of gestures that WUW supports is borrowed from stylus-based interfaces. WUW lets the user draw icons or symbols in the air using the movement of the index finger and recognizes those symbols as interaction instructions. For example, drawing a star (Figure 3G) can launch the weather application. Drawing a magnifying glass symbol takes the user to the map application or drawing an '@' symbol (Figure 3G) lets the user check his mail. The user can undo an operation by moving his index finger forming an 'X' symbol. The WUW system also allows for customization of gestures or addition of new gestures.

WUW prototype

The WUW prototype is comprised of three main hardware components: a pocket projector (3M MPro110), a camera (Logitech QuickCam) and a laptop computer. The software for the WUW prototype is developed on a Microsoft Windows platform using C#, WPF and openCV. The cost of the WUW prototype system, sans the laptop computer, runs about $350.

Both the projector and the camera are mounted on a hat and connected to the laptop computer in the user's backpack (see Figure 1). The software program processes the video stream data captured by the camera and tracks the locations of the color markers at the tip of the user's fingers using simple computer-vision techniques. The prototype system uses plastic colored-markers (e.g. caps of whiteboard markers) as

the visual tracking fiducials. The movements and arrangements of these color markers are interpreted into gestures that act as interaction mechanism for the application interfaces projected by the projector. The current prototype only tracks 4 fingers – index fingers and thumbs. This is largely due to the relative importance of the index finger and thumb in natural hand gestures, tracking these WUW is able to recognize a wide variety of gestures and postures. The maximum number of tracked fingers is only constrained by the number of unique fiducials, thus the WUW also supports multi-user interaction. The prototype system also implements a few object augmentation applications where the camera recognizes and tracks physical objects such as a newspaper, a coffee cup, or a book and the projector projects relevant visual information on them. WUW uses computer-vision techniques in order to track objects and align projected visual information by matching pre-printed static color markers (or patterns) on these objects with the markers projected by the projector.

Discussion and Future Work

In this paper, we presented WUW, a wearable gestural interface that attempts to free digital information from its confines by augmenting surfaces, walls and physical objects and turning them into interfaces for viewing and interacting with the digital information using intuitive hand gestures. We presented the design and implementation of the first WUW prototype which uses a hat mounted projector-camera configuration and simple computer-vision based techniques to track the user's hand gestures.

WUW proposes an untethered, mobile and intuitive wearable interface that supports the inclusion of relevant information into the user's physical environment and vice versa. The object augmentation feature of the WUW system enables many interesting scenarios where real life physical objects can be augmented or can be used as interfaces. WUW also introduces a few novel gestural interactions (e.g. taking photos by using a 'framing' gesture) that present multi-touch based systems lack. Moreover, WUW uses commercially available, off-the-shelf components to obtain an exceptionally compact, self contained form factor that is inexpensive and easy to rebuild. WUW's wearable camera-projector and gestural interaction configuration eliminates the requirement of physical interaction with any hardware. As a result, the hardware components of WUW can be miniaturized without suffering from usability concerns, which plague most computing devices that rely on touch-screens, QWERTY keyboards, or pointing devices.

We foresee several improvements to WUW. First, we plan to investigate more sophisticated computer-vision based techniques for gesture recognition that do not require the user to wear color markers. This will improve the usability of the WUW system. We also plan to improve the recognition and tracking of physical objects and to make projection of interfaces onto those objects and surfaces self-correcting and object aware. We are also implementing another form factor of the WUW prototype that can be worn as a pendant. We feel that if the WUW concept could be incorporated into such a form factor, it would be socially more acceptable, while also improving recognition thanks to the fact that the body of the user moves less than the head. While detailed user evaluation of the WUW prototype has not yet been performed, informal user feedback collected over the past two months has been

very positive. We plan to conduct a comprehensive user evaluation study in the near future.

Acknowledgements

We thank fellow MIT Media Lab members who have contributed ideas and time to this research. In particular, we thank Prof. Ramesh Raskar, Sajid Sadi and Doug Fritz for their valuable feedback and comments.

References

[1] Agarwal, A., Izadi, S., Chandraker, M. and Blake, A. High precision multi-touch sensing on surfaces using overhead cameras. IEEE International Workshop on Horizontal Interactive Human-Computer System, 2007.

[2] Apple iPhone. http://www.apple.com/iphone.

[3] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., MacIntyre, B. Recent Advanced in Augmented Reality. IEEE Computer Graphics and Applications, Nov-Dec. 2001.

[4] Buxton, B. Multi-touch systems that i have known and loved. http://www.billbuxton.com/multitouchOverview.html.

[5] Dietz, P. and Leigh, D. DiamondTouch: a multi-user touch technology. *In Proc. UIST 2001*, Orlando, FL, USA.

[6] Han, J. Low-cost multi-touch sensing through frustrated total internal reflection. *In Proc. UIST 2005*.

[7] Letessier, J. and Bérard, F. Visual tracking of bare fingers for interactive surfaces. *In Proc. UIST 2004*, Santa Fe, NM, USA.

[8] LM3LABS Ubiq'window. http://www.ubiqwindow.jp

[9] Malik, S. and Laszlo, J. Visual touchpad: a two-handed gestural input device. *In Proc. ICMI 2004*, State College, PA, USA.

[10] Matsushita, N. and Rekimoto, J. HoloWall: designing a finger, hand, body, and object sensitive wall. *In Proc. UIST 1997*, Banff, Alberta, Canada

[11] Microsoft Surface. http://www.surface.com

[12] Oblong G-speak. http://www.oblong.com

[13] Perceptive Pixel. http://www.perceptivepixel.com

[14] Pinhanez, C. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. *In Proc. Ubicomp 2001*, Atlanta, Georgia, USA

[15] Raskar, R., Baar, J.V., Beardsley, P., Willwacher, T., Rao, S., Forlines, C. iLamps: geometrically aware and self-configuring projectors, *In Proc. ACM SIGGRAPH 2003.*

[16] Rekimoto, J., SmartSkin: An Infrastructure for Freehand Manipulation on Interactive Surfaces. *In Proc. CHI 2002*, ACM Press (2002), 113-120.

[17] Segen, J. and Kumar, S. Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera. *In Proc. CVPR1999*, (1999), 479-485.

[18] Starner, T., Weaver, J. and Pentland, A. A Wearable Computer Based American Sign Language Recognizer. In P*roc. ISWC'1997*, (1997), 130–137.

[19] Starner, T., Auxier, J., Ashbrook, D. and Gandy, M. The Gesture Pendant: A Self-illuminating, Wearable, Infrared Computer Vision System for Home Automation Control and Medical Monitoring. *In Proc. ISWC'2000*.

[20] Ukita, N. and Kidode, M. Wearable virtual tablet: fingertip drawing on a portable plane-object using an active-infrared camera, *In Proc. IUI2004*, Funchal, Madeira, Portugal.

[21] Wilson, A. PlayAnywhere: a compact interactive tabletop projection-vision system, *In Proc.UIST2005*, Seattle, USA.

[22] Wilson, A. TouchLight: an imaging touch screen and display for gesture-based interaction, *In Proc. ICMI 2004*, State College, PA, USA.