3-D Interaction with a Large Wall Display using Transparent Markers

Akito Hyakutake
hyaku@vogue.is.uec.ac.jp

Koichiro Ozaki
ozaki@vogue.is.uec.ac.jp

Kris Makoto Kitani
kitani@is.uec.ac.jp

Hideki Koike
koike@is.uec.ac.jp

Graduate School of Information Systems
University of Electro-Communications
1-5-1 Chofugaoka Chofu Tokyo 182-8585, Japan

ABSTRACT

In this paper we proposed a new interface for interacting with large displays via small video devices such as a cell phone. We estimate the location of the camera relative to the display using a matrix of transparent markers embedded on display. As a result, our interface allows the user to interact with digital contents without being distracted by opaque visual markers. Our interface enables intuitive interactions such as pointing, rotating, dragging and dropping. Moreover, our use of a small hand-held camera device allows for interaction with large scale displays without the need for direct contact with the display surface. Thus our system is well suited for interactions when there is some distance between the user and the display. Our proposed system has applications to large scale advertisement displays and can enable interactions between individuals and large scale digital content.

Categories and Subject Descriptors

H5.1 [Information Interfaces and Presentation (e.g., HCI)]: : Multimedia Information Systems; Artificial, augmented, and virtual realities

General Terms

Interaction, transparent markers

Keywords

Augmented Reality, Polarization, LCD, Marker, Wall Display, Vision-Based HCI

1. INTRODUCTION

Signs and advertisements (information displays) are ubiquitous in cosmopolitan cities, ranging both in size and location. In the past wall mounted displays (e.g. posters) only conveyed information without any interaction capabilities. However, due to recent advances in touch panel technology and computer vision technology, we often see interactive wall displays where people can actively interact with information (e.g. [5]). When people find interesting information on a display, they can obtain detailed information by utilizing the interaction capabilities which are provided by the display interface.

A representative example of such interactive systems are touch panel based systems (e.g. [9]). However, one of the issues with large touch panel displays is that in a public environment, it requires people to touch the screen. When the display is very large or is mounted at a high position, there are regions of the display that cannot be used interaction.

Computer vision based systems enable non-contact interaction from a distance by recognizing people's faces, gaze direction, finger position, or hand position (e.g. [7]). However, there are only a few systems which are robust enough to work in public environment since vision-based systems are often sensitive to changes in lighting conditions.

One of the solutions to this issue is to use barcodes or visual markers for robust recognition. In Japan, many advertisements make use of two dimensional barcodes called QR code [4], which allow a curious observer to access more information on the web via their QR equipped cellphones. However, a drawback of QR codes is that they introduce meaningless (from a observers perspective) clutter into the display design.

From the above discussion, the requirements for the large wall displays in public environment can be summarized as follows.

1. multiuser interaction
2. non-contact interaction
3. robust recognition
4. interaction at any part of the display

This paper describes a interaction method for large wall-sized displays by using a matrix of transparent 2-D markers [8]. The main contributions of our approach are that (1) it enables multiuser, non-contact, robust interaction without spoiling background images by using transparent markers; (2) it enables accurate localization of pointing position and orientation from a locally captured image by using a matrix transparent markers.

2. RELATED WORK

Ballagas proposed the Point & Shoot [2] system which allows a user to interact with digital content via cellphone camera and an LCD display. However, their systems requires that visible markers are displayed for a duration of time to allow the camera to acquire its position.

The HIEI Projector [10] is a special projector that projects both visible and invisible (infa-red) light onto a screen. The system projects invisible markers by using the infa-red projector and allows for seamless integration of markers and digital content. However, the system presupposes a infa-red source (in this case a projector) and a camera with an infa-red sensor.

The Anoto-pen [1] is another interaction device that makes use of infa-red light. The systems uses a specially designed infa-red pen and encoded paper, to locate the exact position of the pen on the paper. The paper is encoded by very small unique dot patterns, which are recognized by the infa-red sensor attached to the end of the pen, thereby allowing the system to track the exact position of the pen. However, a potential limitation of the system is that it requires the pen to be in contact with the paper. Furthermore, the Anoto-pen is not well suited for interaction at a distance, as was possible with the previous two systems.

In CHI2009, Koike et al[8] presented transparent 2-D markers on an LCD tabletop. Their markers are almost transparent to the human eye but clearly visible as a black-and-white pattern to the camera attached with a polarization sheet. Our method makes use of their transparent markers.

3. SYSTEM OVERVIEW

Our proposed system use a compact hand held camera (e.g. camera equipped mobile phones) to allow for seamless interaction with digital content on large LCD displays (Figure 1). The implementation only requires that a thin transparent marker sheet (polarizer) be placed on the LCD display and on the lens of the imaging device. With this setup the camera device is able to detect the markers to obtain the cameras absolute position (3D translation and rotation) with respect to the display. Knowing the 3D position of the camera allows the camera to be used a pointing device to interact with specific content on the display. Since the markers are invisible to the human eye, the markers do not introduce any kind of clutter to the digital content. Moreover, as long as the markers are visible to the camera, the system allows for interaction over a wide range of distances.

By using camera equipped mobile phones, users can privately access their desired information prompted by the display via the internet. Furthermore, the system can also be used to send information to the display and be shared with other users, to create a public BBS.

4. THE TECHNOLOGY

4.1 Polarization

When light passes through a polarization sheet, it changes the polarization of the light but is still visible to the human eye. Using this fact to our advantage, we fabricate a transparent 2-D barcode marker (Figure 2, 3) on a polarization sheet and attach another polarization sheet on the lens of the camera.

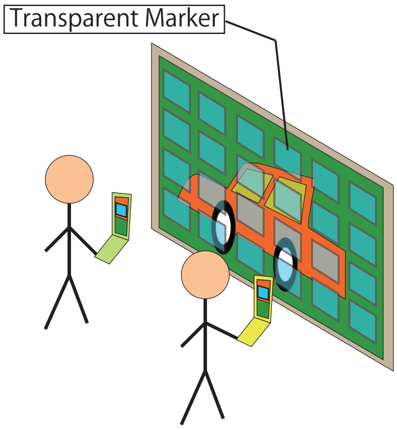


Figure 1 Multiple users interacting with a large LCD display

Linearly polarized light can be transformed to and from circularly polarized light using a quarter-wave polarization sheet. In our implementation we use a combination of quarter-wave polarization sheets to allow light from selected regions of the display to reach the camera (Figure 4).

Figure 2 No polarizer

Figure 3 With polarizer

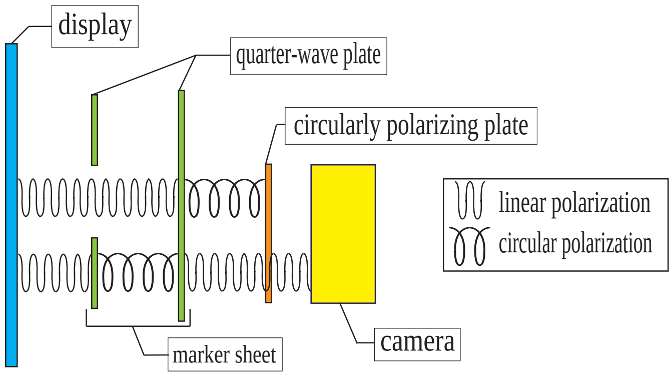
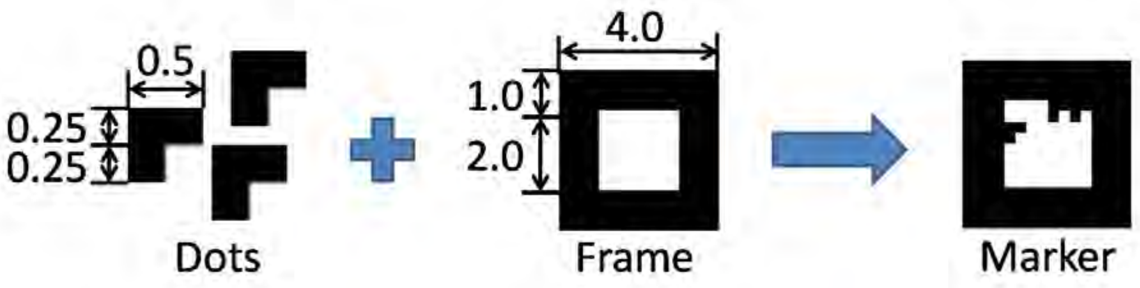


Figure 4 Polarization of light in our system

4.2 Marker Design

Each marker is constructed from a frame and a binary pattern of our basic symbol. The basic symbol of our marker is a rotated “L” shaped binary pattern as seen in Figure 5(a). This symbol allows the camera to recognize the position of the camera. A maximum of 16 (4×4) symbols can be placed within the frame, which yields a total of 2^{16} unique markers.



(a) (b) (c)

Figure 5 Marker design

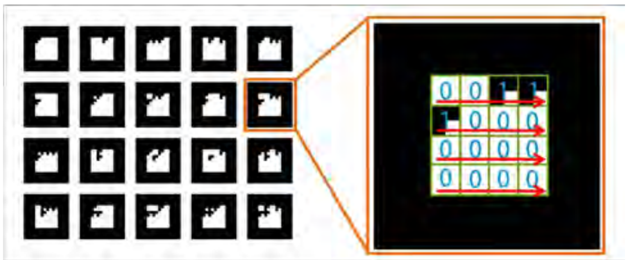


Figure 6 Marker matrix

The aspect ratio of the marker is 1 : 1 as depicted in Figure 5(b) and the absolute size of the marker is assumed to be known by the camera device. By computing the transformation between the detected marker and the actual marker data, we can estimate the relative pose of the camera device.

4.3 Pose estimation

We use the AR toolkit[3] multi-marker recognition algorithm to estimate the 3D pose of the camera device. As long as at least one marker is visible to the camera device the algorithm is able to compute the transformation between the camera image coordinates and the marker coordinates (Figure 7). The marker matrix can be scaled to larger displays by simply increasing the number of markers in the marker matrix.

Given the transformation$\mathbf{T}_{rm}$ which converts the individual marker coordinates to the standard marker matrix coordinate and the transformation $\mathbf{T}_{cr}$ which converts the marker matrix coordinates to the camera coordinates, a point in the marker coordinate space $(x_m, y_m, z_m)^T$ can be mapped to a point in the camera coordinate space $(x_c, y_c, z_c)^T$ using the following transformation.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{T}_{cr}\mathbf{T}_{rm} \begin{bmatrix} x_m \\ x_m \\ x_m \end{bmatrix} \quad (1)$$

The transformation matrix $\mathbf{T}_{cr}$ can be factorized to compute the 3D translation and rotation parameters of the camera device, which will be useful for the application we introduce in the next section.

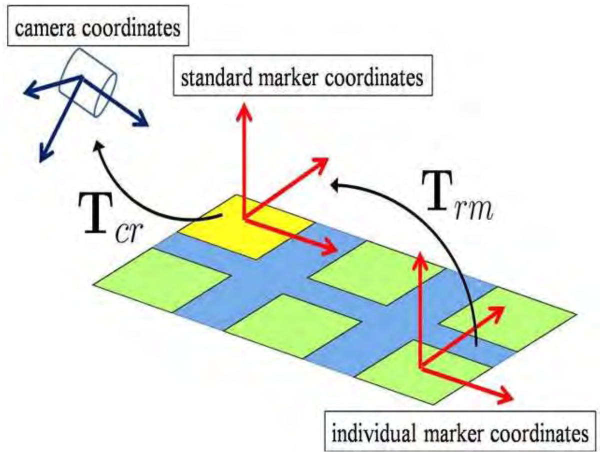


Figure 7 Coordinate transformations

5. INTERACTION EXAMPLES

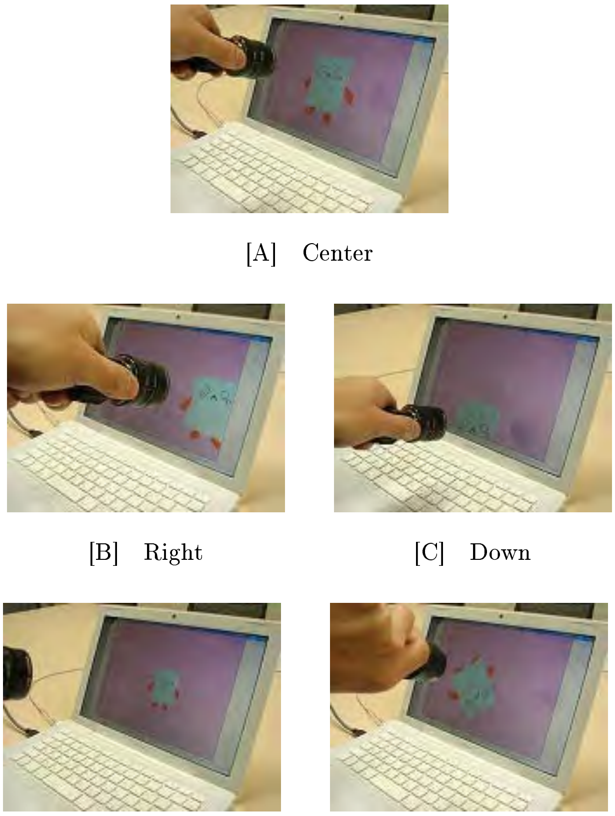
Figure 8 is an example of showing the possible interactions 3D objects displayed on the LCD using a hand held device.

When the user aligns the cameras line of sight to an object on the display, it is possible to calculate the change in 3-D translation and rotation of the device from that position using coordinate system spanned by the invisible markers.

When the user rotates the device while pushing a button, the object can be rotated by calculating the rotation angle of the device relative to the marker matrix. When the user moves the device along the depth axis (distance) to the screen, the size of the object is magnified or reduced, again based on the distance from marker matrix. The system is also robust to translations relative to the screen. Under this setup when the use drags an object along the horizontal (or vertical) axis, the position (2D projection of the camera center on the display) of the object is translated smoothly across the screen by dynamically acquiring the position of the markers in it current view.

6. DISCUSSION

One of the issues of our current implementation is faulty marker recognition for dark background images. Since the "white" regions of the transparent marker allows the light emitted by the display to reach the camera, the marker



[A] Center

[B] Right [C] Down

[D] Far [E] Rotation

Figure 8 Interactions with a 3D-object

recognition fails when the background is very dark. A possible solution is to display a white background for a very short during at a specified frequency (e.g. every 1/60 seconds) and to use a high-frame rate camera. In the past, such high speed cameras were very expensive but recently digital cameras and even some mobile phones are currently equipped with high frame rates.

Another issue with our current implementation is that it is difficult to see the actual displayed image via the handheld device because the markers occlude the image. A possible solution is to use an electronically activated polarizing filter. The polarizing filter can be deactivated when the user is viewing the digital image via the camera and can then be activated again when the user pushes the capture button. Another possible solution is to use two cameras, where one camera is used to capture the background image and the other camera is used to capture the transparent marker. This is viable solution since certain lines of mobile phones are already equipped with multiple cameras.

7. CONCLUSION

We have proposed a system for interacting with wall-sized LCDs with a camera equipped mobile phone, by using a transparent marker matrix. The main contributions of our approach are the following:

1. robust marker based recognition without spoiling the background image;
2. enables accurate estimation of the pointing position by using the transparent marker matrix.

We are currently experimenting with the minimal marker size for camera-based devices. We are also exploring the use of mixed resolution markers to make the system more robust to interactions over a wide range of distances from the screen.

8. REFERENCES

[1] Anoto Group AB. Anoto pen. http://www.anoto.com/

[2] R. Ballagas, M. Rohs, J. G. Sheridan, Sweep and Point & Shoot: Phonecam-Based Interactions for Large Public Displays. In Conference on Human Factors in Computing System: CHI'05 extended abstracts on Human factors in computing systems, Vol. 2, pages 1200-1203, 2005

[3] H. Kato and M. Billinghurst. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99), pages 85-94, 1999.

[4] Denso Corp. QR Code. http://en.wikipedia.org/wiki/QR_Code/

[5] GestureTek, GestPoint. http://www.gesturetek.com/

[6] Hecht, Eugen, Optics, 4th Edition, published by Pearson Education, Inc., publishing as Benjamin Cummings, Copyright 2002 , pages 163-172. 2009.

[7] H. Koike, M. Toyoura, M. Ebata, K. Oka, Y. Sato. Human Interaction with a Wall-Sized Display using a Stereo Camera, The 7th International Conference on Ubiquitous Computing (UbiComp 2005), CD-ROM proceedings, P42, 2005.

[8] H. Koike, W. Nishikawa, K. Fukuchi. Transparent 2-D Markers on an LCD Tabletop System. In Proceedings of the 27th international conference on Human factors in computing systems(CHI'09), pages 163-172. 2009

[9] Peltonen, Peter and Kurvinen, Esko and Salovaara, Antti and Jacucci, Giulio and Ilmonen, Tommi and Evans, John and Oulasvirta, Antti and Saarikko, Petri, It's Mine, Don't Touch!: interactions at a large multi-touch display in a city centre, Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems (CHI 2008), pages 1285–1294, 2008.

[10] Y. Shirai, M. Matsushita, T. Ohguro, HIEI Projector: Augmenting a real environment with invisible information In The 11th Workshop on Interactive Systems and Software (WISS2003), pp.115-122,2003