Presentation assistant and kiosk interaction with fiducial markers

Crease, Murray; Fiala, Mark

National Research Council Canada
Conseil national de recherches Canada

Institute for Information Technology
Institut de technologie de l'information

NRC · CNRC

*Presentation Assistant and Kiosk Interaction with Fiducial Markers **

Crease, M., Fiala, M.
October 2007

* published at the IEEE International Workshop on Haptic Audio Visual Environments and their Applications (HAVE'2007). October 12-14, 2007. Ottawa, Ontario, Canada. NRC 49876.

Canada

HAVE 2007 - IEEE International Workshop on
Haptic Audio Visual Environments and their Applications
Ottawa, Canada, 12-14 October 2007

Presentation Assistant and Kiosk Interaction with Fiducial Markers

Murray Crease, Mark Fiala
People-Centred Technologies, Institute for Information Technology
National Research Council Canada
46 Dineen Drive
Fredericton, NB, E3B 9W4
{murray.crease,mark.fiala}@nrc-cnrc.gc.ca

Abstract – *Fiducial markers and computer vision software provide a low cost and flexible way for human-computer interaction. Using the high reliability and minimal image pixel requirements of the* ARTag *fiducial marker system we are able to identify users and perform control actions with a simple small printed pattern. We describe two novel applications. Firstly, a user giving a presentation can control the slides with simple motions of a small printed marker towards their laptop or portable computer's built-in camera, thus replacing a remote or the necessity to step back towards the computer to advance the slides. Secondly, patterns printed on a small card can be used to interact with a kiosk, an example is the back of a ID badge used to bring up the schedule and direction information to a conference attendee.*

Keywords – *HCI, User Interface, Fiducial Marker, ARTag.*

I. INTRODUCTION

The use of mouse and keyboard to interact with Graphical User Interfaces (GUIs) is largely effective for desktop computers. Mobile devices such as cell phones and Personal Desktop Assistants (PDAs) have become increasingly common and afford users the freedom to interacting with information systems while mobile. A third paradigm is the emerging trend of pervasive computing systems where users are able to interact with publicly available displays without keyboards, mice or handheld devices. Gesture recognition and touch screens are examples of interaction techniques where users are able to approach a display and interact with no specialised technology and training. Vogel and Balakrishnan suggest that such systems may make mobile devices redundant [5]. The authors describe four possible zones of interaction with such displays: ambient - the user is simply walking past; implicit - the user turns his head towards the display; subtle - the user approaches the display; and personal - the user interacts directly with the display.

Computer vision had not yet provided complete solutions for robust gesture recognition in uncontrolled settings and identity recognition. However, marker-based computer vision has achieved an appropriate level of reliability for these tasks. We present two systems where it is not inconvenient to use such markers.

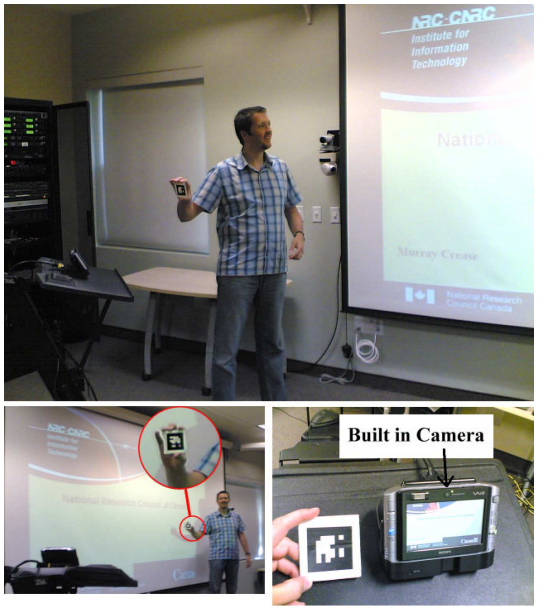

Fig. 1. *Fiducial marker pattern used to control presentation slides.*

A. Using Fiducial Marker Systems for HCI

A fiducial marker system is designed to solve the following problem: given an input image (either a static image or a frame from a video stream) provide a list of markers found in the image. This list is defined as consisting of a list of points, each point is described by its image location (preferably with sub-pixel resolution) and a label such as an ID number. The ARTag system [1], [2] provides a list of **4n** points of sub-pixel precision where each point is labeled as to which of the four corners of an ARTag marker it belongs, and what the ID number of the marker is.

This information can be used in different ways, the simplest being to use only the marker ID to trigger some behavior dependent only on the presence of a marker such as logging into a kiosk. In [4] a scuba diver signals instructions to an underwater robot by holding up cards with ARTag markers. The image position information can also be used, such as using a

marker as a hand-held computer mouse pointer as in the work of Longworth [6]. The center, or just one of the corner locations, is utilized for this. More complex applications such as augmented reality can use all four corners to calculate a 6 degree of freedom pose. Complex 3D objects can have markers covering them, where the full pose is calculated in every camera frame to give 3D input device functionality [3].

We describe two prototype systems herein which use a single marker, and use simpler interaction modes to replace a remote control to enable a user to control the slide deck of their presentation, and for a user to bring up information at a kiosk.

While the presence of the fiducial markers and their ID numbers are reliably detected, and their position and orientation within the image found with little noise (typically less than 0.1 pixels of position jitter for a marker center), the depth or rotation about axes other than the main camera axis is less accurate [6]. Thus we designed two simple interation techniques using the strengths of: 1) presence detection, 2) reliable ID, and 3) rotation detection within the camera plane. The remote control replacement uses the marker presence detection and rotation strengths to allow a user to advance or rewind the slides in their presentation (typically Microsoft Powerpoint). The user flashes a small card, such as two inches wide, towards their portable computer on the podium and makes a clockwise or counterclockwise rotation movement to advance or rewind their slides. We used the portable Sony Vaio UX-280P tablet PC which has a built in 640x480 video camera. The camera has a field of view of 30 degrees, which coupled with the minimum pixel size of 17 pixel of the ARTag markers means the two inch pattern can detected up to 8 feet away, a reasonable marker size and operation distance for a typical presenter. A benefit of this system is that, other than the obvious of not requiring batteries, it can be easily replaced if lost by simply reprinting the pattern unlike existing remote devices. The pattern could be printed on the back of a business card for convenience. The system has been used in several presentations and is shown in Fig. 1.

The second application relies simply on the strengths of: 1) presence detection and 2) reliable ID to allow control of a kiosk. In our envisaged application, visitors to a conference could have an ARTag marker printed on the back of their ID card, they can show this to the kiosk and it will bring up the information relevant to them. For example, a map could have their poster location highlighted, the time and location of a presenation they are to give, and perhaps display some recent messages from other conference attendees or staff. The ARTag system allows 1001 different markers which should suffice for smaller to medium conferences.

II. CONCLUSIONS

Two simple systems utilizing a fiducial marker system were created for a conference environment: controlling a presentation and operating a kiosk. The strengths of the ARTag passive vision based fiducial marker system relevant to an HCI were identified, that of accurate recognition, identification, and rotational pose. Future work includes using the position to allow more complex input commands to kiosk devices.

Fig. 2. *A conference attendee using their ID badge to access information from a kiosk.*

[1] M. Fiala, "ARTag, a Fiducial Marker System Using Digital Techniques", Proc. *IEEE Intl. Computer Vision and Pattern Recognition Conference 2005 (CVPR'2005)*, pp. 590-596, 2005.

[2] M. Fiala, "ARTag Revision 1, a Fiducial Marker System Using Digital Techniques.", Proc. *IEEE Intl. Computer Vision and Pattern Recognition Conference 2005 (CVPR'2005)*, National Research Council Publication 47419/ERB-1117, Nov, 2004.

[3] M. Fiala, "3D Input Using Hand-held Objects and Computer Vision", Proc. *IEEE Intl. Workshop on Haptic Audio Visual Environments and their Applications*, pp. 94-98, 2006.

[4] G. Dudek and J. Sattar and A. Xu, "A Visual Language for Robot Control and Programming: A Human-Interface Study", Proc. *IEEE Intl. Conference on Robotics and Automation(ICRA)*, pp. 2507-2513, Apr 2007.

[5] D. Vogel and R. Balakrishnan, "Interactive Public Ambient Displays: Transitioning from Implicit to Explicit, Public to Personal, Interaction with Multiple Users", Proc. *17th ACM Symposium on User Interface Software and Technology (UIST 2004)*, pp. 137-146, October 2004.

[6] R. Longworth, "Interaction with Large Screen Displays using ARTag Marker System", *Senior Thesis Report, Faculty of Computer Science, University of New Brunswick, 2007*, March 2007.