User-Following Displays

Gopal Pingali, Claudio Pinhanez, Tony Levas, Rick Kjeldsen, Mark Podlaseck
IBM T.J. Watson Research Center
19 Skyline Drive, Hawthorne, NY 10532 USA
gpingali | pinhanez | levas | fcmk | podlasec@us.ibm.com

Abstract

Traditionally, a user has positioned himself/herself to be in front of a display in order to access information from it. In this information age, life at work and even at home is often confined to be in front of a display device that is the source of information or entertainment. This paper introduces another paradigm where the display follows the user rather than the user being tied to the display. We demonstrate how steerable projection and people tracking can be combined to achieve a display that automatically follows the user.

1. Introduction

Displays are a dominant aspect of our life at work and at home today, with the proliferation of computers, televisions, game consoles and other information and entertainment devices. Current office design, work-style, and even living room design are driven largely by the need to be in front of a fixed display. How will our style of functioning and the design of our working and living spaces change if displays are not tied to specific devices or surfaces? Is it possible for a display to appear on an appropriate surface depending on where the user is? In this paper, we introduce this notion of a user-following display and present one method to realize such a display.

We propose steerable projection as a way to emancipate a display from fixed devices or surfaces. Figure 1 shows a device called the Everywhere Displays projector (ED-projector) [6] which integrates an LCD projector and a computer-controlled mirror to steer a display on to any surface in an environment. As an image is steered on to different surfaces, it has to be pre-corrected for oblique projection distortion, as discussed in [6]. As seen in Figure 1, a steerable camera is also combined with the steerable projector. The camera can be used to recognize a user's hand gestures and to allow user interaction with the projected display as discussed in [4]. While previous papers [6,4] have presented the steerable projection and vision-based interaction aspects of the ED-projector, this paper focuses on how a user-following display can be achieved with the ED-projector.

We define a user-following display as one that appears on an appropriate surface in an environment, at the appropriate size, shape, and orientation, depending on the position of a user in the environment. While there are other interesting aspects such as adapting the resolution and colors of the display, depending on the user's position and the current display surface, we restrict ourselves here to the issue of appropriately positioning the display.

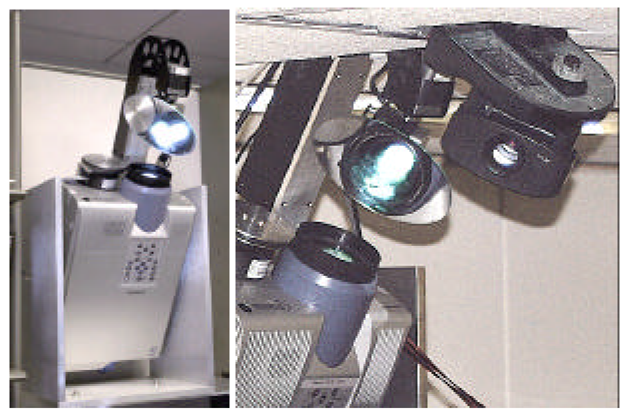

Figure 1. Everywhere Displays projector: a) projector with rotating mirror; b) detail of a complete prototype with an integrated pan/tilt/zoom camera.

There has been significant amount of previous research in augmenting reality and creating large-scale displays using projected images [7, 9, 12]. However, the issue of a user-following display has not been addressed. The work that comes closest is [8] where cameras are used to track a user and a computer that is close to the user is automatically activated. Our work is unique in combining steerable projection with people tracking. Here, the user-following display can appear on virtually any surface.

A user-following display system based on the ED-projector should: (i) have a model of the environment (especially, all the displayable surfaces); (ii) track the position and orientation of the user in the environment; (iii) dynamically select a display surface based on the position of the user; (iv) determine the orientation, shape, and size of a display zone on the selected surface; (v) choose the appropriate projection parameters for the selected display zone (including focus and zoom of the projector, orientation of the mirror, and the image pre-warping transformation); and (vi) project the image using these parameters. Importantly, the selection of the display zone should ensure that the user does not occlude the steered projected display.

There are several challenges in realizing such an automatic user-following display. Some of these include model update, discovery of display surfaces, dynamic calibration/state update of the projector-mirror system, and determination of head orientation and field of view of a user. While we continue to work on these issues, in this paper we present a first version of a user-following display that relies on pre-calibration of display surfaces/zones.

2. A Pre-calibration Based Method for User-Following Displays

Figure 2 shows an overview of our proposed method for user-following displays. A camera-based tracking system dynamically determines the position of a user in an environment. Multiple cameras are used to cover a large environment. We use camera calibration parameters [10] to map the image location of a user to world coordinates. A search for an appropriate display zone is then performed over a set of pre-specified display zones. A pre-calibration procedure for each display zone yields the projection parameters and the homography [1] that warps an image so that its projection is aligned with the display zone. Once a display zone is dynamically selected based on user tracking, the corresponding parameters drive the projection system resulting in a display that follows the user.

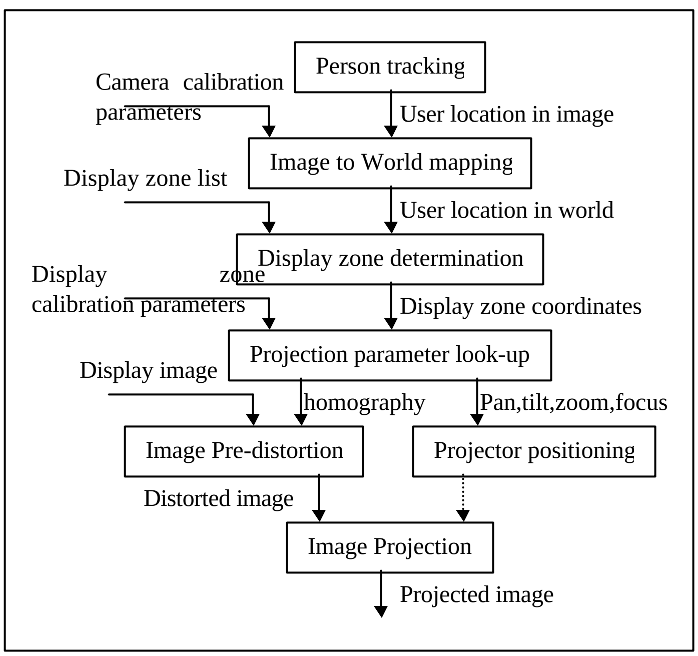

Figure 2. Overview of pre-calibration based method for a user-following display.

2.1 Environment modeling and pre-calibration

The user-following display system maintains a 3D model of its environment. Figure 3 shows an example VRML model of a lab environment, built using measurements made in the lab. Possible display zones in the environment are identified beforehand. Figure 3 shows several pre-specified display zones. The system stores the world coordinates for each display zone. When a user specifies a display zone, he or she also performs a projector calibration procedure. Currently, we provide the user with a graphical interface to control the orientation of the ED-projector's mirror, and the zoom and focus of the projector. The user projects a rectangular calibration pattern on the display zone. This pattern is not aligned with the display zone due to oblique distortion. The user then uses the graphical interface to align the vertices of the projected rectangle with the vertices of the display zone. In this paper, we assume quadrilateral planar display zones.

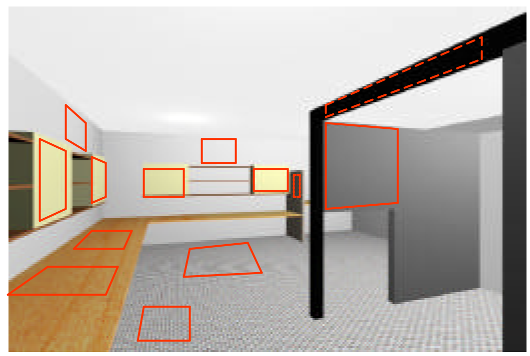

Figure 3. A 3D environment model with specified display zones (dashed lines indicate a display zone on the opposite face of the object).

The mapping between the original image coordinates of the rectangular pattern and the image coordinates after alignment with the display zone specify the homography for pre-warping any image to appear in the display zone. After this calibration procedure, the system stores the homography parameters as well as the focus, zoom, and orientation parameters of the ED-projector corresponding to the specified display zone.

2.2 User Tracking

In this version of the user-following display system, we adopt a real-time camera-based head tracking technique to determine the position of the user in the environment. The head tracking technique is based on motion, shape, and flesh-tone cues. Similar head tracking approaches have been reported in [3,5,11]. We first perform a differencing operation on consecutive frames of the incoming video and threshold the result. A morphological closing operation then removes noise and fills up small gaps in the detected motion regions. A standard contour-tracing algorithm then yields the bounding contours of the segmented regions. We then smooth the contours and compute the orientation and curvature along the contour.

We then analyze the shape of each contour to check if it could be a head. We look for curvature changes corresponding to a head-neck silhouette (concavities at the neck points and convexity at the top of the head) with the head pointing up and for the circularity of a head shape. We then check for sufficient flesh-tone color within the detected head region by matching the color of each pixel within the head contour with a model of flesh tone colors in normalized r-g space. This technique detects multiple heads in real time. In the current system, we track a single user in the environment (see Figure 4).

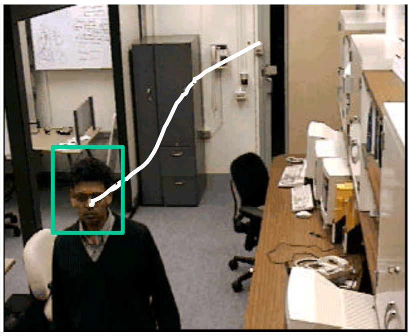

Figure 4. Illustration of camera-based head tracking

Once the head region is detected in an image, we estimate the corresponding position of the head in world coordinates by making several assumptions. We assume that the user is upright and that the approximate height of the user is known. Using camera calibration parameters derived from the procedure in [10], we find the intersection of a line from the optical center through the detected centroid of the head in the image plane with a horizontal plane at the assumed height of the user's head.

2.3 Display Zone Selection

After estimating the user location in an environment, we select a display zone by searching over the set of pre-determined display zones (see Section 2.1), based on some constraints. The display zone selection policy depends on the application, and is typically driven by knowledge of the position and orientation of the user. We have implemented a simple policy in this version of the system. The center of a zone should be at least 2 feet from the head location and no more than 12 feet away. The line from the head location to the center of a display zone should be less than 45° from the normal to the display zone. This line should also be less than 45 degrees in tilt and 90 degrees in pan from the viewing direction of a user. Within these constraints, the display zone that is closest to the user is selected.

While we are working on estimating head orientation, a user's viewing direction is not directly estimated in our current system. The system maintains a history of head tracking information and assumes the user's angle is aligned with the direction of motion. The constraint on the tilt angle is enforced assuming a horizontal viewing direction for the user. Determining a user's head orientation or gaze is an active area of research (see, [2] for example).

2.4 Occlusion detection

An important consideration is that the user should not be in the way of the projection for the selected display zone. Therefore, after selecting a display zone as described in Section 2.3, the system performs an occlusion check. To do this, the projector has to be calibrated in 3D. We perform such calibration as an additional step during the display zone calibration procedure described in Section 2.1. This procedure is very similar to the camera calibration procedure described in [10]. In camera calibration, a set of known points in the real world are matched with their corresponding locations in the camera's image plane. The projector calibration procedure records the real-world coordinates of a set of projected points with their known locations in the projector's image plane. The computation of projector parameters is then identical to the computation of camera parameters in [10].

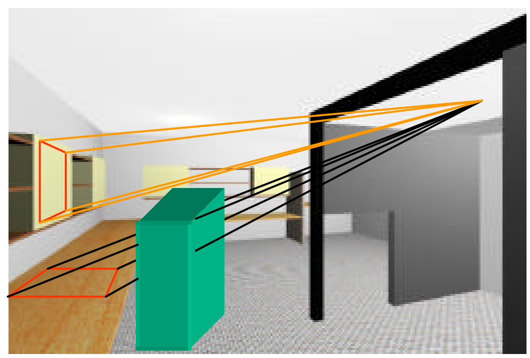

Figure 5. Detection of occlusion of a display zone. In this situation, the occlusion check results in the selection of the upper display zone over the lower one.

To detect occlusion, we check for intersection between the frustum from the optical center of the projector to the vertices of the display zone and a box defining the bounding volume of a user (see Figure 5). If occlusion is detected, the next best display zone is selected (see Section 2.3).

3. Results

We implemented a user-following display system in a 24'x20' lab with a height of 9 feet. A single ED-projector is placed towards the middle of the room closer to one of the longer walls. From this position, the ED-projector is able to hit most of the surfaces in the room except on the wall behind it. We use two cameras, each tracking the user over a different part of the room. We defined more than 30 display zones in this environment on various surfaces such as desks, filing cabinets, walls, whiteboards, and the floor.

Figure 6 shows examples from a sequence where the display (picture of a cartoon face) follows the user as he moves around the environment. The image to the left of the figure indicates the path taken by the user (in red). The upper part of the image shows the camera that tracks the user as he walks along this path. During the sequence, the display moves across seven different zones on the walls, the four filing cabinets, and the desk. Four of these are shown in the images to the right of Figure 6, corresponding to the marked positions on the path.

One of the limiting factors in our system is that the ED-projector's mirror takes about 5 seconds to precisely switch between surfaces. We are currently able to rapidly move the display across different zones on a large single surface such as a wall. Here, we keep the mirror static and zoom the projector out to cover the whole surface.

4. Conclusions

We have presented the concept of a user-following display and realized such a display using steerable projection and camera-based tracking. The method presented here involves a significant calibration overhead. However, this one-time overhead is acceptable and practical in a number of application environments.

Several steps can be taken to improve the effectiveness of the system presented here. These include precise mirror motion to ensure continuous steering between surfaces, tracking with stereo cameras to more accurately determine the 3D location of the user, and tracking of the orientation of the head. A significant research challenge is to build a system that performs dynamic self-calibration and display zone selection instead of relying on pre-calibrated display zones.

References

1. Faugeras, O., Three-Dimensional Computer Vision: A Geometric Viewpoint. 1993, Cambridge, Massachusetts: The MIT Press.
2. Gee, A. and Cipolla, R., Determining the gaze of faces in images. Image and Vision Computing, vol. 12, no. 10, p. 639-647, 1994.
3. Gibbon, D. and J. Segen, Detecting people in image sequences for electronic camera panning applications. Technical Report BL0113520-930201-03TM, AT&T Bell Laboratories, 1993.
4. Kjeldsen et al. Interacting with Steerable Projected Displays. To apppear in 5th International Conference on Automatic Face and Gesture Recognition. 2002. Washington, DC.
5. Pingali, G., Tunali, G. and Carlbom, I. Audio-Visual Tracking for Natural Interactivity. In Proc. of ACM Multimedia. 1999. Orlando, Florida.
6. Pinhanez, C. The Everywhere Displays Projector: A Device to Create Ubiquitous Graphical Interfaces. In Proc. of Ubicomp'01. 2001. Atlanta, Georgia.
7. Raskar, R. Oblique Projector Rendering on Planar Surfaces for a Tracked User. In Proc. of SIGGRAPH'99. 1999. Los Angeles, California.
8. Shafer et al. The Easy Living Intelligent Environment System. In Proc. of CHI Workshop on Research Directions in Situated Computing, April 2000.
9. Sukthankar, R., et al. Smarter Presentations: Exploiting Homography in Camera-Projector Systems. In Proc. of ICCV'01. 2001. Vancouver, Canada.
10. Tsai, R.Y. An efficient and accurate camera calibration technique for 3D machine vision. In Proc. of CVPR, 1986, pp. 364-374.
11. Wang, C. and Brandstein, M., Multi-source Face Tracking with Audio and Visual Data . In Proc. of IEEE International Workshop on Multimedia Signal Processing, Copenhagen, Denmark, 1999.
12. Welch, G., et al., Projected Imagery in Your "Office of the Future". IEEE Computer Graphics and Applications, 2000(July/August): p. 62-67.

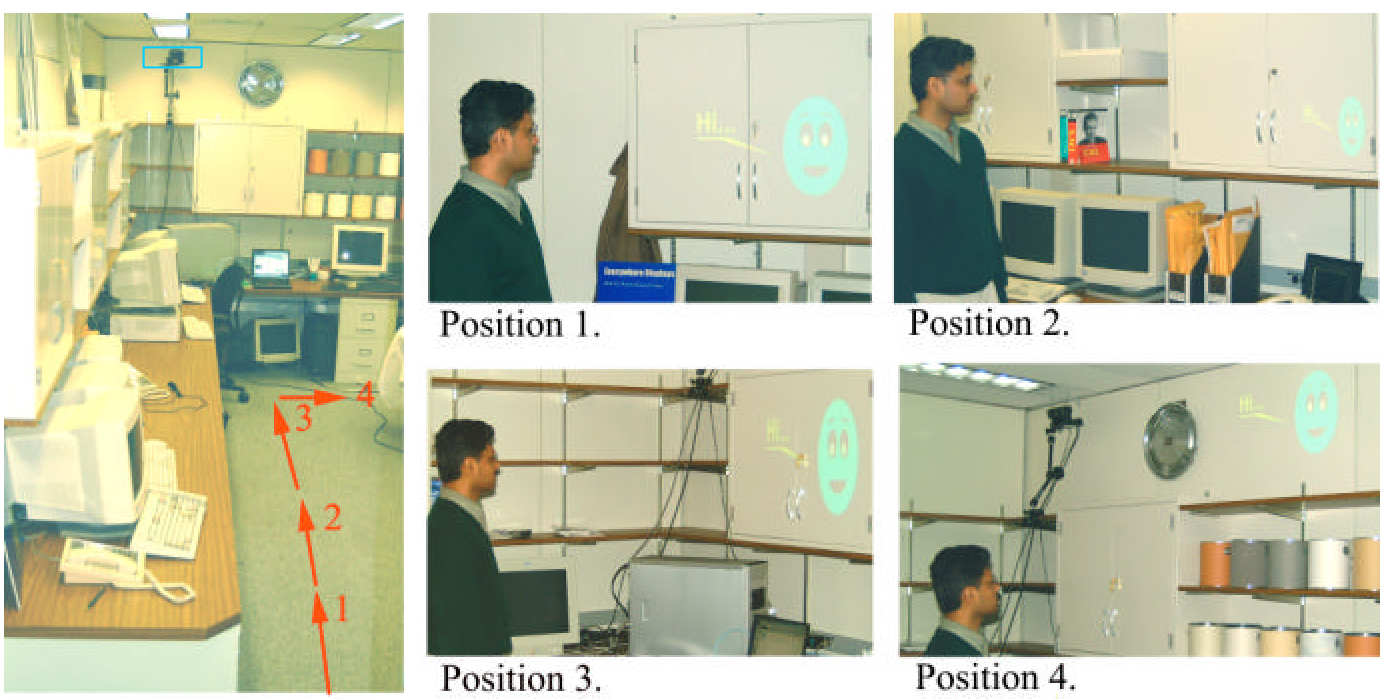

Fig. 6. Example sequence illustrating the user-following display. Left: An image of the environment showing the path followed by a user. The tracking camera is seen at the top. Right: Four images showing how the display moves to different zones depending on the user's position (corresponding user positions are marked in the image on the right).