A HANDHELD MIRROR SIMULATION

Alexandre R.J. François and Eun-Young Elaine Kang
Integrated Media Systems Center
University of Southern California
Los Angeles, CA 90089
{afrancoi,ekang}@usc.edu

ABSTRACT

We present the design and construction of a handheld mirror simulation device. The perception of the world reflected through a mirror depends on the viewer's position with respect to the mirror and the 3-D geometry of the world. In order to simulate a real mirror on a computer screen, images of the observed world, consistent with the viewer's position, must be synthesized and displayed in real-time. Our system is build around a LCD screen manipulated by the user, a single camera fixed on the screen, and a tracking device. The continuous input video stream and tracker data is used to synthesize, in real-time, a continuous video stream displayed on the LCD screen. The synthesized video stream is a close approximation of what the user would see on the screen surface if it were a real mirror. Our system provides a generic interface for applications involving rich, first-person interaction, such as the Virtual Daguerreotype.

Keywords
Human-Computer Interface; Real-time video processing.

1.INTRODUCTION

We present the design and construction of a handheld mirror simulation device. Rather than attempting to solve the difficult vision problem of a general, exact mirror simution, our focus is on the design and construction of a first person Human-Computer Interface (HCI) to be used in new interactive experiments and applications. While remaining simple, our first prototype must provide a better experience to the user than the traditional fixed camera / fixed display / horizontal image flip mirror setting. A major novelty in our approach is the hands-on aspect, which requires a more elaborate model to convincingly reproduce the interaction between the user and the mirror as an object. One important clue is the evolution of the user's field of view in the mirror as her position with respect to the mirror changes. Furthermore, our goal is to produce a working but extensible system, to serve as a platform for future enhancements to the HCI device subsystem itself, as well as to allow easy integration in applications

Our system is composed of a handheld LCD screen manipulated by the user, a single camera fixed on the screen, and two 6 degrees of freedom trackers giving in real time the positions and orientations of the camera and the user's head. The major system components and their relations are illustrated in figure 1. The user sees on the screen what they would see on a real mirror.

The perception of the world reflected through a mirror depends on the viewer's position with respect to the mirror and the 3-D geometry of the world. In order to simulate a real mirror on a computer screen, images of the observed world, consistent with the viewer's position, must be synthesized and displayed in real-time. The continuous input video stream and tracker data is used to synthesize, in real-time, a continuous video stream displayed on the LCD screen. The synthesized video stream is a close approximation of what the user would see on the screen surface if it were a real mirror.

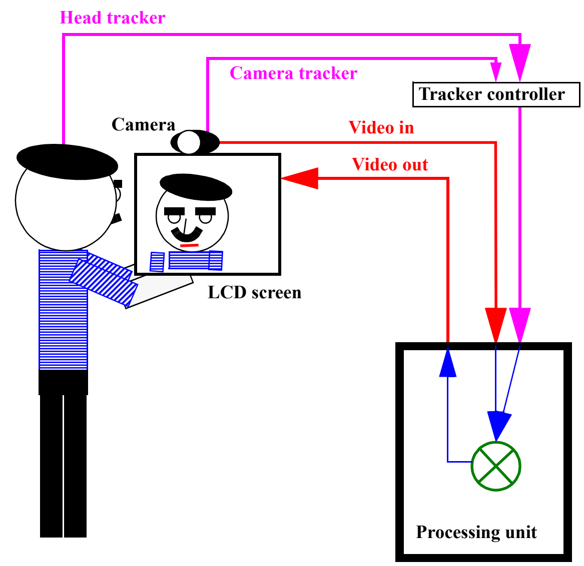

Figure 1. Handheld mirror simulation system.

This paper is organized as follows. In section 2, we outline the geometric analysis of mirror and camera imaging, leading to the image transform for mirror simulation, based on the flat world assumption. In section 3, we briefly explain the software and hardware design of our prototype. In section 4, we relate our work to other mirror experiments, and to upcoming applications. Finally, we conclude the paper in section 5.

2.GEOMETRIC ANALYSIS

2.1 Mirror geometry

When a person looks into a planar mirror, they see what they would see from the symmetrical (virtual) point of view (see figure 2). Assimilating image formation by the eye to a perspective projection, the image perceived by the viewer on the mirror surface is

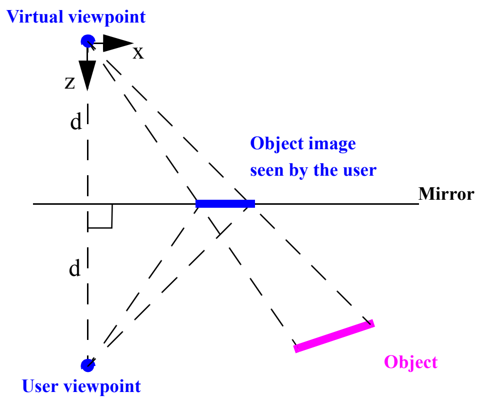

Figure 2. Mirror geometry.

a perspective projection on the mirror plane with respect to the virtual viewpoint. A point of homogenous coordinates $[xyz1]^T$ in an orthonormal coordinate system centered on the virtual viewpoint (symmetrical of the user viewpoint with respect to the mirror plane) with the z direction normal to the mirror, is seen on the mirror surface at its projected location $[x_p y_p d]^T$, where d is the orthogonal distance of the viewpoint to the mirror plane, and

$$\begin{bmatrix} U_p \\ V_p \\ W_p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/d & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \qquad \begin{matrix} x_p = U_p/W_p = dx/z \\ y_p = V_p/W_p = dy/z \end{matrix}$$

The image perceived on the mirror depends only on the position of the user with respect to the mirror, and the field of view covered by the mirror depends only on the position of the user with respect to the mirror edges. In particular, the image perceived on the mirror remains "fixed" when the viewing direction changes while the viewpoint remains fixed.

2.2 Pinhole camera geometry

A camera whose position and orientation are fixed with respect to the mirror produces a video stream that corresponds to this particular viewpoint. Assuming the camera can be approximated by a pinhole model (see figure 3), a point of homogenous coordinates

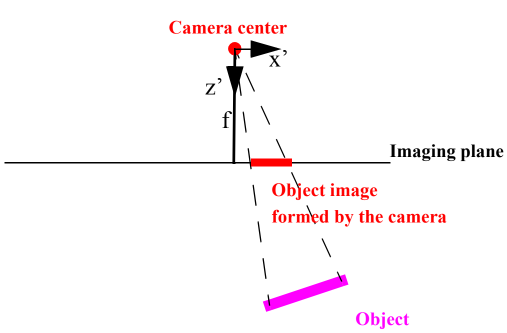

Figure 3. Pinhole camera geometry.

$[x'y'z'1]^T$ in an orthonormal coordinate system centered on the camera center, is projected on the imaging plane at the location $[x'_p y'_p f]^T$, where f is the focal length (orthogonal distance of the camera center to the imaging plane), and

$$\begin{bmatrix} U'_p \\ V'_p \\ W'_p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \qquad \begin{matrix} x'_p = U'_p/W'_p = fx'/z' \\ y'_p = V'_p/W'_p = fy'/z' \end{matrix}$$

2.3 User view synthesis

In order to synthesize the images corresponding to the user's viewpoint from the images captured by the camera, we must relate them geometrically, i.e. express x_p and y_p in terms of x'_p and y'_p.

We place the camera so that the image plane and (virtual) mirror planes coincide. We chose the viewpoint and camera coordinate systems so that their respective axes directions coincide. Thus the 4x4 matrix relating the two coordinate systems is reduced to a simple translation of the origin:

$$\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix}$$

The mirror projection equations become:

$$\begin{bmatrix} U_p \\ V_p \\ W_p \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1/d & t_z/d \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} \qquad \begin{matrix} x_p = d(x' + t_x)/(z' + t_z) \\ y_p = d(y' + t_y)/(z' + t_z) \end{matrix}$$

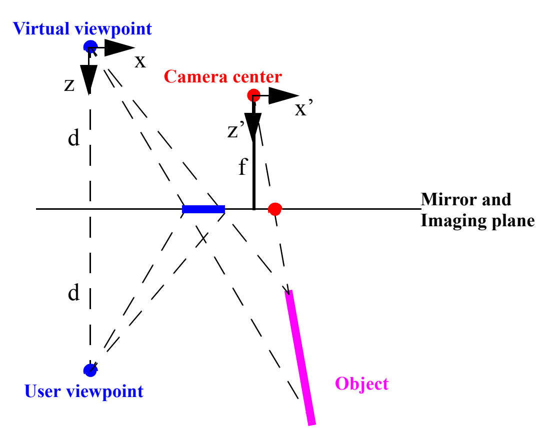

Figure 4. An underconstrained problem.

Express x_p and y_p in terms of x'_p and y'_p thus requires to recover the actual 3-D position of each point from its image projection. This is obviously an underconstrained problem, as an image pixel might potentially be the projection of any world point on a ray passing through the camera center and the pixel, as illustrated in figure 4.

Various computer vision techniques could be used to attempt to recover the 3-D geometry of the world from images in some cases. Another solution would be to use a 3-D model of the environment, if available. For our first prototype, we followed a simpler approximate approach. Since we are designing our mirror as a first person interaction tool, we assume that the center of interest is the user himself, and that the background is of lesser importance. We thus model the world as a plane parallel to the mirror/imaging plane, at the same distance as the user's viewpoint (see figure 5).

The z' component of each world point is thus fixed to $z' = f + d$ in the camera centered coordinate system. We can now inverse the camera projection equations:

$$x'_p = fx'/(f+d) \Leftrightarrow x' = (f+d)x'_p/f$$
$$y'_p = fy'/(f+d) \Leftrightarrow y' = (f+d)y'_p/f$$

The transformation equations that relate the user's view to the camera image are thus:

$$x_p = d((f+d)x'_p/f + t_x)/((f+d) + t_z)$$
$$y_p = d((f+d)y'_p/f + t_y)/((f+d) + t_z)$$

The only parameters are f, d and [txtytz]. The theoretical focal length of the camera can be obtained by various calibration methods. The positions and orientations of the user's viewpoint and the camera center provided by the tracking devices, allow to compute both the distance d and the change of referential translation parameters.

Note that under our assumptions, the transform relating the user's view to the camera view is reduced to a scaling and a translation. This model provides a straightforward and efficient way of synthesizing the viewer's image stream from the camera stream, while remaining a reasonable approximation for a large class of applications. The major field-of-view variations occuring when varying the angle of view and the distance to the mirror, are convincingly simulated.

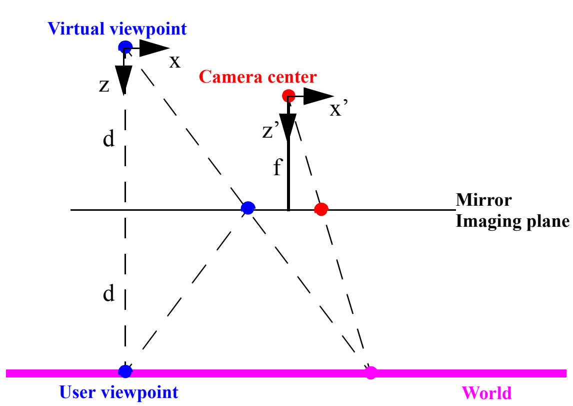

Figure 5. Flat world assumption.

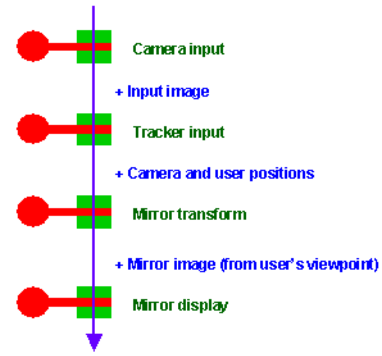

Figure 6. Mirror simulation application graph.

3.SYSTEM INTEGRATION

Beyond establishing a simple image transform to simulate reflexion in a mirror, a central goal of our work is to build an efficient and extensible working system.

We have implemented the mirror image transform as a module in the Modular Flow Scheduling Middleware [2]], the open source implementation of IMSC's software architecture for Integrated Media Systems. It is specifically designed for parallel processing of concurrent data streams in real-time [3], and its modularity dramatically reduces development and integration time.

The application graph for the mirror simulation is presented in figure 6. Generic modules such as video capture and display were already available. Note that thanks to the modularity of the framework, any processing can be carried on the images before or after the mirror transform.

The prototype system is composed of a handheld 10" LCD screen fitted with a camera and a magnetic tracker sensor (see figure 7), a fixed processing unit, a fixed magnetic tracker transmitter, and a second magnetic tracker sensor placed on the user. Real-time camera and tracker inputs are used to synthesize the mirror image stream and display it in real-time on the LCD screen. The result is a convincing handheld mirror simulation, in spite of the inherent limitations of our mirror transform model.

4.RELATED WORK AND APPLICATIONS

Mirrors have been source of inspiration for artists for a long time, but few interactive electronic mirror experiments have actually been demonstrated. The "Magic Morphin Mirror" [1] uses vision-based face tracking to present a distorted appearance of the viewer's reflection. "Wooden Mirror 1999" [5], and more recent experiments by its creator are other examples of electronic mirror

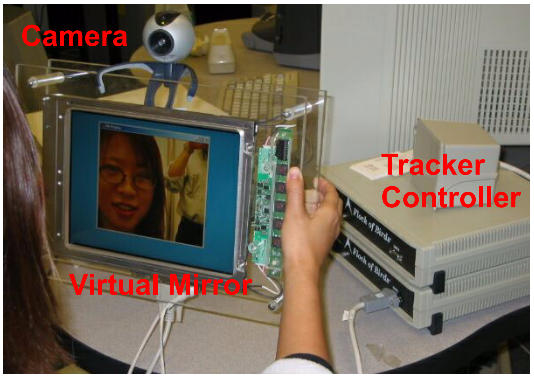

Figure 7. Handheld mirror device.

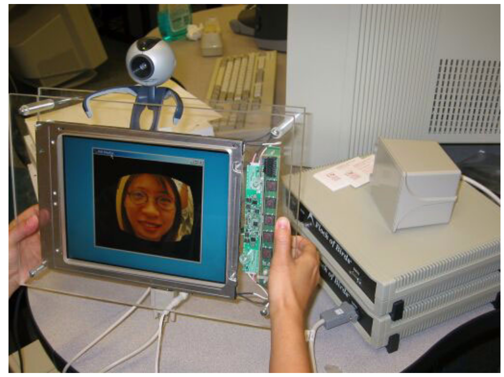

Figure 8. Concave mirror simulation.

metaphors. Our project aims to open the door to even more powerful experiments by providing a realistic, geometrically sound mirror simulation. A major innovation in our approach is to allow and encourage the user to freely move around, and interact with the mirror physically, although other mirror settings, such as a large (simulated) mirror on a wall, are currently being designed.

Our system provides a fundamental platform for other real-time interactive applications. Digital image processing modules can easily be added to the system in order to simulate various mirror deformations, such as magnifying, concave and convex mirrors (see figure 8). Note that the deformation parameters can also be dynamic.

We are currently using the mirror simulation platform for the Virtual Daguerreotype project [4]. Daguerreotypes are one of the earliest forms of photography, a direct-positive process that creates a highly detailed image on a sheet of copper plated with a thin coat of silver without the use of a negative. As daguerreotype images, most of which are portraits, are encased in velvet, leather, glass and metal trim, and require viewing from a certain angle, they seem to belong in one's hand. However, the fragility of these objects demands that they be displayed in cases, under carefully controlled light. By compositing pictures of daguerreotypes, taken from different angles, with the synthetic mirror simulation images, our system will allow to recreate the essence of looking into a mirror and seeing a changing engraved image superimposed on one's reflection in real time.

5.CONCLUSION

We have presented a handheld mirror interface system. The system is designed to be a platform that will be used in many interactive applications. In particular, the use of the Modular Flow Scheduling Middleware as the software architecture ensures efficiency and extensibility through multithreading and modularity.

Adopting the simplifying pinhole camera and flat world models, we have shown that the transform relating the camera view and the user view depends only on the relative positions of the user viewpoint and the camera center, and is reduced to a scaling and a translation. This model provides a straightforward and efficient way of synthesizing the viewer's image stream from the camera stream, and provides a reasonable approximation in a large class of applications. Beyond the inherent inability of the model to account for depth in the scene, a major limitation is the impossibility to simulate eye contact. This is due to the physical separation between the capture device (camera) and the visualization device (screen). More elaborate modeling and processing is required to achieve this effect, e.g. with head modeling, tracking and synthesis

The design and construction of realistic mirror simulations represents an original experiment in Human-Computer Interface, that opens the door to various interactive experiments in entertainment, arts and communication, such as the Virtual Daguerreotype. Use of video analysis and graphics techniques will allow to explore and interfere with what has always been a private, solitary act of a person looking in a mirror.

6.ACKNOWLEDGMENTS

The authors wish to thank M. Lazzari, M. McLaughlin, Gérard G. Medioni and W.L. Wong at the University of Southern California. This research has been funded by the Integrated Media Systems Center, a National Science Foundation Engineering Research Center, Cooperative Agreement No. EEC-9529152; and by a grant from the USC Arts Initiative.

7.REFERENCES

[1] Darrell T. et al., "Magic Morphin Mirror: Face-Sensitive Distortion and Exaggeration," *SIGGRAPH'97 Electric Garden, Visual Proceedings*, p.120, 1997.

[2] François, A. R.J. *Modular Flow Scheduling Middleware*. http://mfsm.sourceForge.net/

[3] François, A. R.J. and Medioni, G. G. "A Modular Software Architecture for Real-Time Video Processing." *Proc. International Workshop on Computer Vision Systems*, pp. 35-49, Vancouver, Canada, July 2001.

[4] Lazzari, M., McLaughlin, M. L., Jaskowiak, J., Wong, W., and Akbarian, M. A haptic exhibition of daguerreotype cases for USC's Fisher Gallery. In *Touch in Virtual Environments: Haptics and the Design of Interactive Systems*, McLaughlin, M. L., Hespanha, J., and Sukhatme, G., Eds., IMSC Series in Multimedia, Prentice-Hall.

[5] Rozin D., "Wooden Mirror 1999," *SIGGRAPH 2000 Electronic Art Exhibit*, p.68, 2000. Online: http://www.itp.nyu.edu/danny