Olwal, A. and Hollerer, T. **POLAR: Portable, Optical see-through, Low-cost Augmented Reality**. *Proceedings of VRST 2005 (ACM Symposium on Virtual Reality and Software Technology)*, Monterey, CA, Nov 7-9, 2005, pp. 227-230.

POLAR: Portable, Optical see-through, Low-cost Augmented Reality

Alex Olwal[1, 2] Tobias Höllerer[1]

[1]University of California, Santa Barbara
Department of Computer Science
Santa Barbara, CA 93109
USA

[2]Royal Institute of Technology (KTH)
Department of Numerical Analysis and Computer Science
100 44 Stockholm
Sweden

alx@kth.se, holl@cs.ucsb.edu

ABSTRACT
We describe POLAR, a portable, optical see-through, low-cost augmented reality system, which allows a user to see annotated views of small to medium-sized physical objects in an unencumbered way. No display or tracking equipment needs to be worn. We describe the system design, including a hybrid IR/vision head-tracking solution, and present examples of simple augmented scenes. POLAR's compactness could allow it to be used as a lightweight and portable PC peripheral for providing mobile users with on-demand AR access in field work.

Categories and Subject Descriptors
H.5.1 [Multimedia Information Systems]: Artificial, augmented, and virtual realities; H.5.2. [User Interfaces]: Graphical user interfaces, Input devices and strategies; I.3.6 [Methodology and Techniques]: Interaction Techniques.

General Terms
Human Factors

Keywords
Augmented reality, optical see-through, portable, compact, low-cost, projection.

1. INTRODUCTION
Augmented Reality (AR) merges virtual information with the view of a real environment. It presents an intuitive and direct user interface to annotate physical objects without physically altering them. AR has proven itself as a useful paradigm in many different application areas, including medicine, maintenance and repair, architecture, and entertainment [1]. Most AR systems require encumbering equipment, such as head-worn displays, active or passive stereo glasses, and position and orientation trackers to produce the registered augmentation effect. Also, with the notable exclusion of mobile AR [7], AR environments typically consist of large stationary setups.

Figure 1. The POLAR system. The prototype AR setup uses a laptop, a custom-made foldable mount, a beam splitter (half-silvered mirror), a camera and an IR rangefinder.

We are interested in enabling portable and unencumbered AR that is spontaneously available without incurring large over-head costs in either setup time or price of equipment. Many field-work based applications could profit from AR views onto the physical objects they are dealing with. The limited availability of stationary AR environments, combined with general inconvenience of setup and usage, hinder easy exploration of AR as a tool in application areas such as archaeology, geology, and botany. As a first step towards more ubiquitous unencumbered AR, we present a low-cost portable AR "chamber" that can visually annotate small to medium-sized objects such as machine parts, archaeological artifacts or geological specimens, while not requiring the user to wear any equipment. Our optical see-through system, POLAR, uses a standard laptop display as the image source, a half-silvered mirror as the optical combiner and a camera and IR sensor to track the user's viewpoint (see Figure 1 and 2).

In the remainder of this paper we discuss related work in Section 2, followed by an overview of our system in Section 3. Section 4 and 5 describe our hybrid tracking components and results. Finally, conclusions and future work are presented in Section 6.

Olwal, A. and Hollerer, T. **POLAR: Portable, Optical see-through, Low-cost Augmented Reality**. *Proceedings of VRST 2005 (ACM Symposium on Virtual Reality and Software Technology)*, Monterey, CA, Nov 7-9, 2005, pp. 227-230.

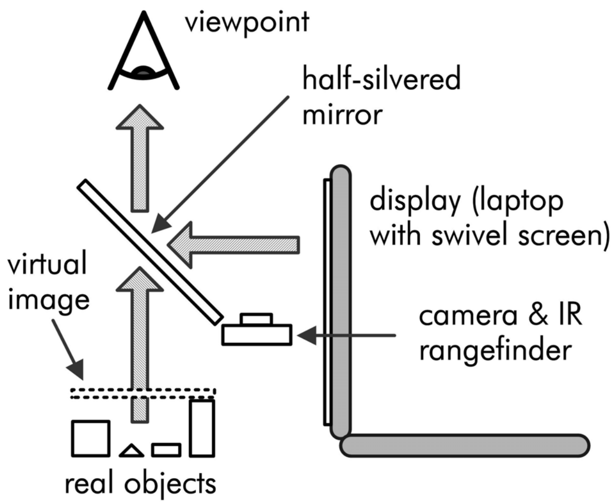

Figure 2. Overview of the POLAR system in a downward-looking configuration. The graphics generated on the laptop are reflected on the half-silvered mirror, which simultaneously allows the user to see the real-world objects below it. The camera and IR rangefinder track the user's viewpoint.

2. RELATED WORK

Optical see-through head-worn displays (HWDs), which date back to Sutherland's pioneering work in the 1960s [16], often employ half-silvered mirrors as beam-splitters. While HWDs are a convenient way of creating AR views onto an environment, we are interested in AR solutions without encumbering equipment. Other AR systems project the graphics directly onto real-world objects [13] and there is a large body of work on tangible user interfaces that annotate paper documents and other physical objects on desk surfaces onto which the computer-generated image is projected [17], [9]. However, this restricts the location of virtual information to physical surfaces, whereas our goal is to be able to place information anywhere in 3D space (POLAR is currently limited to a specified, but adjustable, projection plane).

Previous projection-based optical see-through systems [2] do not meet our goals in terms of portability, ready availability of components, and low cost. Our proposed AR system is related in spirit to the sonic flashlight [15] which overlays ultrasound imagery directly on top of the examined body part. Such domain-dependent AR tools can provide good usability by concentrating on serving a specific application. However, we are pursuing a more general-purpose AR setup. Our suggested solution is a hybrid between a truly mobile AR system [7] and a stationary setup, in that we designed the system to have a minimal footprint and easy setup.

The unencumbered AR experience distinguishes our approach from setups that employ HWDs and sensors on the user. Additionally, POLAR's use of optical see-through provides a direct view of the real objects, in contrast to video see-through displays, such as those used in mobile AR applications [5].

3. THE POLAR SYSTEM

Our prototype provides a small workspace where real objects are overlaid with virtual information. Similarly to most optical see-through systems (e.g., [2]), virtual information is created by reflecting a display image onto a half-silvered mirror. As shown in Figure 2, we are using a standard laptop display and beam splitter. The image appears to lie in a plane behind the mirror and we can adjust the location of the 2D plane by varying the distance of the laptop display, in contrast to optical see-through systems that project onto transparent surfaces [6]. Our system can be set up in various configurations and viewing angles. The standard configuration of Figure 1 and 2 has the user look downwards onto a table top where the annotation target is placed.

POLAR's compactness greatly simplifies assembly, disassembly and transportation of the system, as shown in Figure 3.

4. HYBRID TRACKING

To avoid worn equipment and the use of expensive special-purpose hardware (e.g., stereo cameras), we utilize a combination of 2D eye tracking and a distance sensor.

4.1 2D Eye Tracking

We use a commercial SDK [11] to track the user's eyes in a grayscale 320×240 video from a firewire camera [3]. The software tracks facial features in the video stream and was extended to send the data to our application over UDP. While we currently treat the center between the eyes as the eye coordinate, we can also base it on the right eye, since most people are right-eye dominant.

Figure 3. The POLAR system components in folded state. Clockwise from top left: Laptop, external soundcard (allows laptop to connect to the A/D converter's MIDI interface over USB), A/D converter with MIDI interface, IR rangefinder, folded mount with stored half-silvered mirror, and firewire camera. The external sound card and A/D converter will be replaced by a custom USB-based distance sensing module, about the size of the camera.

Olwal, A. and Hollerer, T. **POLAR: Portable, Optical see-through, Low-cost Augmented Reality**. *Proceedings of VRST 2005 (ACM Symposium on Virtual Reality and Software Technology)*, Monterey, CA, Nov 7-9, 2005, pp. 227-230.

4.2 1D Distance Sensing

A Sharp GP2Y0A02YK IR rangefinder [14] determines the distance to the closest object, which in our case is the face (assumed to be the same distance from the rangefinder as the eyes). A commercial A/D converter [8] converts the voltages output by the sensor into a 7-bit value. We derived the non-linear mapping from the 7-bit value to the distance experimentally by measuring the output in 1 cm intervals for a target in a 25-75 cm range. A third-degree polynomial describing the relationship was acquired through least mean squares curve-fitting of the obtained data and this equation is used to extract the actual distance in the Fusion module, as described below.

Similarly to the Eye Tracker, the sensor output from the distance sensing module is sent over UDP to the main application.

4.3 Fusion

Our application listens to known UDP ports for input from the 1D distance sensing and the 2D eye tracker. This does not only allow us to abstract the tracking from the main application, but also makes it possible to offload computationally expensive tracking to a separate computer, if needed.

The streamed eye coordinates and distance allows the viewpoint to be easily recovered using the known image size and horizontal/vertical field of view [3]:

$$x = z \times \tan(\frac{x_{eye}}{width_{image}} \times fov_H)$$

$$y = z \times \tan(\frac{y_{eye}}{height_{image}} \times fov_V)$$

where x_{eye} and y_{eye} are the eye coordinates (pixels), $width_{image}$ and $height_{image}$ the image dimensions (pixels), z the distance (mm), fov_H and fov_V the horizontal/vertical field of view (degrees), and (x, y, z) the viewpoint (mm).

Figure 4 illustrates the straightforward mapping of 2D eye coordinates from the eye tracker to real-world coordinates.

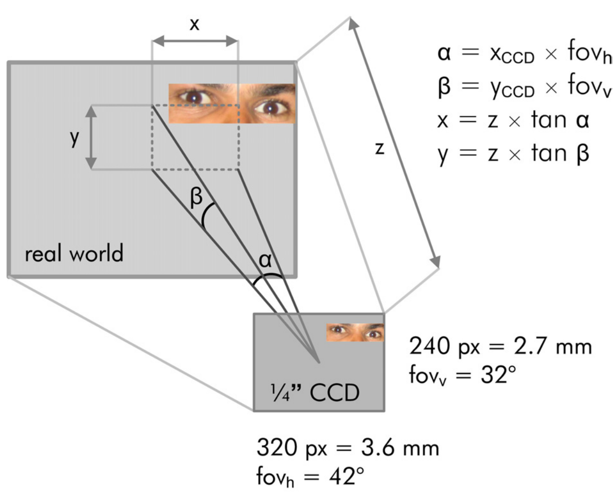

Figure 4. Data from an IR rangefinder (z) and eye tracking software (x,y) allows the viewpoint to be recovered using the known CCD and image characteristics.

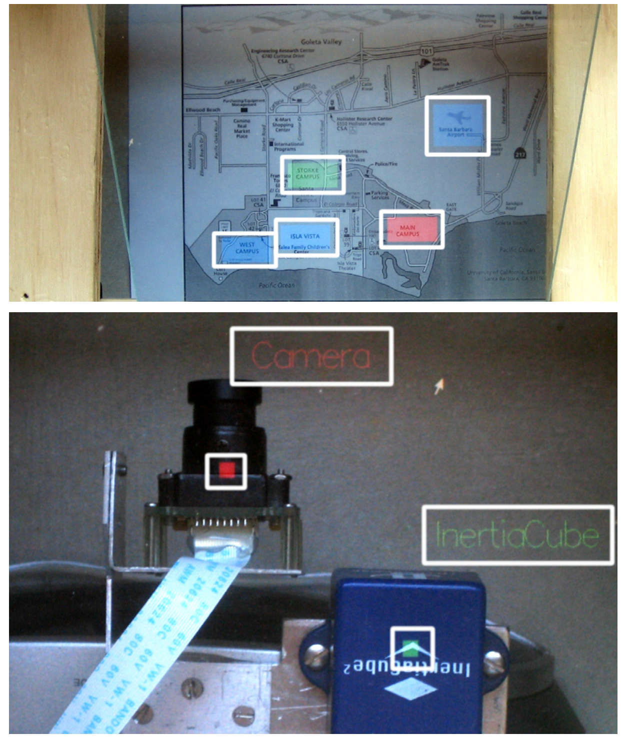

Figure 5. AR view through our display, as the user looks downwards. Top: The AR overlay highlights different locations on a physical map using colored rectangles. Bottom: A camera and a sensor on a piece of equipment are augmented with colored annotations. (The overlays have been marked with white rectangles for clarity.)

4.4 Visualization

The scene is rendered with OpenGL and the user's viewpoint is used to correct the perspective and position of overlaid information as the user's viewpoint changes. Thus as the user will move slightly to the left, the overlaid geometry will follow, such that it seems stable over the superimposed scene.

5. RESULTS

We implemented two proof-of-concept examples within our testbed; an application that highlights locations on a physical map, and an application that provides annotations on top of technical equipment (See Figure 5).

We ran all parts of our system locally (using localhost for sending and receiving UDP data) on a Toshiba Tablet PC, 1.8 GHz, 1.25 GB RAM with a GeForce FX Go5200. We did not experience any performance problems, since the eye tracker is currently the only CPU-intensive module.

The system is manually calibrated by first adjusting the depth of the projection plane and then aligning the overlays with the real-world objects.

The hybrid tracking works well with relatively slow head movement – fast motion will cause the eye tracker to loose track of the feature points, which will result in a few seconds of reinitialization. The eye tracker also depends on having the full face visible in the camera image, with tracking being lost if large parts of the user's face are outside of the image for extended time.

Additionally, due to limitations in current the A/D converter, only 7-bits of precision can be obtained for the distance, which affects the quality of overlay placement.

Our setup is fortunately not dependent on a particular tracking solution – the modular approach provides transparent tracking to the application. It is thus straightforward to replace the current proof-of-concept tracking modules with minimal changes to our system.

6. CONCLUSIONS AND FUTURE WORK

POLAR is a cost-effective, portable, and non-intrusive prototype AR system based on commercial off-the-shelf components and custom software. Our system can be set up quickly and avoids the need for worn trackers through the use of vision-based tracking.

We are addressing our tracking issues by implementing a custom-made distance sensing module which will provide us with 10-bit A/D conversions. One or more SHARP IR rangefinders will be used and the module will connect directly to the host over USB, where the data will be transmitted over UDP using our existing tracking protocol. The implementation of our own eye tracker, with higher resolution (640×480) and color for improved results, will allow us to adapt the tracking to our specific conditions and increase robustness when the user is close to the camera.

The inclusion of tracked real-world objects and haptic feedback will be used to furthermore increase the level of interactivity and immersion. Tracking of the viewed objects can be achieved with marker-based systems such as ARToolkit [10] or ARTag [4]. However, in view of the small working space of our system, it may be more beneficial to develop a simple custom marker tracking scheme based on smaller color markers.

Additionally, we are currently investigating more complex applications, noting that interaction with projection-based optical see-through displays is still an open research problem.

Besides improving the current prototype and creating a larger version, we are also planning to extend the setup for use with an autostereoscopic optical see-through display [12]. Such multi-view displays could also be used to provide several users with a unique view of the scene, overcoming the single user limitation of our current implementation.

We envision low-cost POLAR systems as portable PC peripherals in the future that, given their compactness, can provide mobile users with on-demand AR access in field work.

7. REFERENCES

[1] Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., and MacIntyre, B. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, 25(6):24-35, Nov-Dec 2001.

[2] Bimber, O., Encarnação, L.M., and Schmalstieg, D. Augmented Reality with Back-Projection Systems using Transflective Surfaces. *Proc. EUROGRAPHICS 2000*, vol. 19, no. 3, 2000, pp. 161-168

[3] Fire-I Digital Camera, Technical Specifications. http://www.unibrain.com/support/tech_specs/fire-i_camera/tech_specs.htm. June 2005.

[4] Fiala, M. ARTag, a Fiducial Marker System Using Digital Techniques. *Proc. IEEE CVPR 2005. Computer Vision and Pattern Recognition,* Volume 2, June 20-26 2005, pp. 590 – 596.

[5] Henrysson A. and Ollila M. UMAR - Ubiquitous Mobile Augmented Reality. *Proc. Third International Conference on Mobile and Ubiquitous Multimedia (MUM2004)*, pp 41-45. College Park, Maryland, U.S.A. October 27 - 29, 2004.

[6] Hirakawa, M. and Koike, S. A Collaborative Augmented Reality System using Transparent Display. *Proc. ISMSE'04*, 2004, pp. 410-416.

[7] Höllerer, T. and Feiner, S. Mobile augmented reality. In H. Karimi and A. Hammad, editors, *Telegeoinformatics: Location-Based Computing and Services*. Taylor and Francis Books Ltd., London, UK, 2004.

[8] I-CubeX. http://infusionsystems.com/. June 2005.

[9] Ishii, H. and Ullmer, B. Tangible Bits: Towards Seamless Interfaces between People Bits and Atoms. *Proc. ACM CHI 97.* 1997, pp. 234-241.

[10] Kato, H., Billinghurst, M., Poupyrev, I., Imamoto, K., and Tachibana, K. Virtual object manipulation on a table-top AR environment. *Proc. IEEE ISAR 2000*, 2000, pp. 111-119

[11] NevenVision Facial Feature Tracking SDK. *http://www.nevenvision.com/*. June 2005.

[12] Olwal, A., Lindfors, C., Gustafsson, J., Kjellberg, T., and Mattson, L. ASTOR: An Autostereoscopic Optical See-through Augmented Reality System. *Proc. IEEE ISMAR 2005*, Oct 5-8, 2005.

[13] Raskar, R., Welch, G., Low, K-L., Bandyopadhyay, D. Shader Lamps: Animating Real Objects With Image-Based Illumination, *Proc. Eurographics Workshop on Rendering 2001*, 2001, pp.89-102.

[14] Sharp GP2Y0A02YK IR rangefinder. *http://sharp-world.com/products/device/lineup/opto/sensor/*. June 2005.

[15] Stetten, G. and Chib, V. *Overlaying Ultrasound Images on Direct Vision*. Journal of Ultrasound in Medicine, vol. 20, no. 3, 2001, pp. 235-240.

[16] Sutherland, I. A Head-Mounted Three Dimensional Display. *Proc. FJCC 1968*. 1968, pp 757-764.

[17] Wellner, P. *Interacting with Paper on the DigitalDesk*. Communications of the ACM, 36 (7), 1993, pp. 87-97.