A Testbed for Precise Registration, Natural Occlusion and Interaction in an Augmented Environment Using a Head-Mounted Projective Display (HMPD)

Hong Hua, Chunyu Gao,
Leonard D. Brown, Narendra Ahuja
Beckman Institute
University of Illinois at Urbana-Champaign,
Urbana, IL 61801
Email: Honghua@uiuc.edu

Jannick P. Rolland
School of Optics-CREOL
University of Central Florida
Orlando, FL 32816

Abstract

A head-mounted projective display (HMPD) consists of a pair of miniature projection lenses, beam splitters, and displays mounted on the helmet and retro-reflective sheeting materials placed strategically in the environment. It has been recently proposed as an alternative to existing 3D visualization devices. In this paper, we first briefly review the HMPD technology, including its featured capabilities and the recent development in both display implementations and applications. Then the implementation of a testbed, namely "playing 'GO' game with a remote opponent in a 3D augmented environment," is described. The testbed not only demonstrates the capabilities of virtual-real augmentation and registration, natural occlusion of virtual objects by real, interaction with augmented environments, as well as networking collaboration, but also embodies part of our long-term objective to develop a collaborative framework in 3D augmented environments. Through the testbed, major calibration issues, such as accommodation/convergence considerations and determination of viewing transformations, are studied and discussed in detail. Both calibration methods and results are included, which are applicable to other applications. Finally, experimental results of the testbed implementation are presented.

1. Introduction

See-through HMDs (STHMDs) superimpose virtual objects on an existing scene to enhance rather than replace the real scene and are widely used in augmented reality (AR) domains [1, 2]. Video and optical fusion are the two basic approaches to combining real and virtual images. In a video see-through HMD, the real-world view is captured with two miniature video cameras mounted on the headgear, electronically merged with virtual objects, and then presented as a combined scene to the user through an immersive HMD. The resolution of video cameras is the resolution limit of the real-world view. The major challenges are the generation of photo-realistic synthetic scenes and precise registration. In an optical see-through HMD, the direct view of the real world is maintained and the computer-generated virtual scene is optically superimposed onto the real scene. While video see-though provides the possibility of attenuating the real-world view to match with its virtual counterpart, optical see-through presents the real world in full resolution, real time, and with wide range of luminance. Therefore the challenge of the optical see-through is to make the virtual scene match with its real counterpart in terms of resolution, realism, real time, and luminance. Comparison and analysis of the various visualization technologies can be found in [3].

Though video see-through is very popular and adequate in many applications, video can never be as good as real. Therefore, optical see-through is required in many demanding applications such as medical planning. The conventional optical see-through displays, however, still have open challenges. One of the problems is occlusion contradiction between virtual and real objects. Occlusion, the hiding of more distant objects by closer ones, is one of the important cues to depth perception. Figure 1 exemplifies an occlusion contradiction existing in conventional optical see-through HMDs.

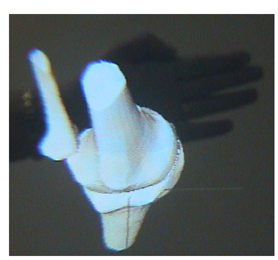


Fig.1 Real object (hand) in front of the virtual object (bone) cannot occlude the virtual object in conventional optical see-through HMDs

The concept of head-mounted projective

Proceedings of the IEEE Virtual Reality 2002 (VR'02)
1087-8270/02 $17.00 © 2002 **IEEE**

displays (HMPDs) was proposed as an alternative to 3D visualization devices [4, 5, 6,7]. Potentially, the HMPD concept can partially provide solutions to the issues existing in the state-of-art visualization devices. For example, the combination of projection and retro-reflection makes it possible to achieve correct occlusion of virtual objects by real objects.

The subject of this paper is to briefly review the HMPD technology and our recent prototype implementation, and to present our testbed for virtual-real augmentation and registration, natural occlusion and interaction, as well as network collaboration using the HMPD technology. Technical implementation, calibration issues, and preliminary results will be included in detail.

2. Head-mounted projective displays

2.1. Overview of the concept

An HMPD, conceptually illustrated in Fig.2, consists of a pair of miniature projection lenses, beam splitters, and displays mounted on the head and a supple and non-distorting retro-reflective sheeting material placed strategically in the environment [4, 5, 7]. An image on the miniature display, which is located beyond the focal point of the lens rather than between the lens and the focal point as in a conventional HMD, is projected through the lens and retro-reflected back to the exit pupil, where the eye can observe the projected image. Because of the retro-reflective property (Fig.3), in which the rays hitting the surface are reflected back on themselves in the opposite directions, the location and size of the perceived image are independent of the location and shape of the retro-reflective screen [7].

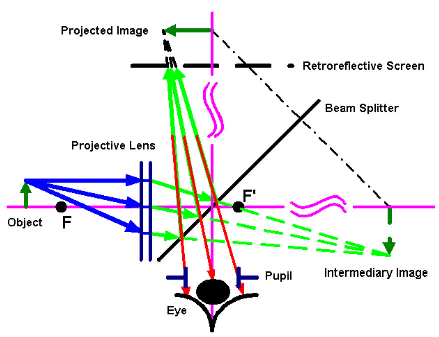


Fig.2 Imaging concept of HMPD

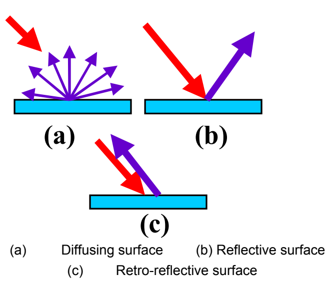


Fig. 3 Behavior of different reflective surfaces.

2.2. Technology highlights

There are two major aspects that distinguish the HMPD technology from conventional HMDs and stereoscopic projection systems, making it appropriate for a wide range of applications: the projective optics, rather than an eyepiece as used in conventional HMDs, and a retro-reflective screen rather than a diffusing screen as used in other projection-based displays [7, 8]. The direct see-through capability allows optically augmenting the real world with computer-generated objects. Furthermore, the HMPD technology provides intrinsically correct occlusion of computer-generated virtual objects by real objects, as demonstrated in Fig. 4, creates a ubiquitous display environment in which a retro-reflective material can be applied anywhere in physical space and can be tailored to arbitrary shapes without introducing additional distortion to the virtual images, and supports unique perspective for each participant in a multiple-user collaborative environment without introducing crosstalk from other participants [7, 8]. Moreover, compared with conventional eyepiece-based optical see-through HMDs (STHMDs), the utilization of projective optics allows for reduced optical distortion across similar FOVs achieved in conventional eyepiece-based HMD designs. Finally, projection optics better meets the pupil size and eye relief requirements [7].

Fig.4 A real object (hand) in front of the virtual object (bone) occludes the virtual object in HMPDs

2.3. Related work

The HMPD concept was initially patented by Fergason in 1997 [4]. Parsons and Rolland [5, 9] have explored its potential medical applications. Hua and Rolland et. al. have made efforts to demonstrate the feasibility of the HMPD imaging concept and to quantify some of the properties and behaviors of the retro-reflective materials in imaging systems [6, 7]. Adopting a Double Gauss lens structure, they built their first-generation prototype from commercially available components [7]. Kawakami and Tachi et al. [10, 11], developed a similar optics configuration named X'tal Vision and proposed the concept of object-oriented display and visual-haptic display. Kijima et. al. [12] did some research and application work. Hua and Rolland et. al. furthered their efforts with an ultra-light (i.e. 8g), high quality projection lens by introducing a diffractive optical element (DOE) as well as plastic components [8, 13], and implemented a compact prototype using the custom-designed lens [8]. The prototype achieves 50 degrees FOV and weighs about 750 grams. Figure 5 (a) shows the front view of the prototype. An image was projected through the display and the picture in Fig. 5(b) was taken at the exit pupil of the optics where the eyes can observe the images. The retro-reflective material is approximately 0.6m or arm-length away from the helmet.

3. A testbed for registration, natural occlusion, and interaction: playing augmented 'GO' game with a remote opponent

Applying the HMPD technology, we have been developing a testbed named "Playing 'GO' game with a remote opponent in a 3D augmented environment" [14]. Through the 'GO' game, our objective is to demonstrate and evaluate the capabilities of virtual-real augmentation, registration, natural occlusion of real/virtual objects, interaction, as well as networking collaboration, which fits well with our long-term objective: to support distance collaboration in 3D augmented environments with intuitive interaction.

3.1. About 'GO'

The equipment to play the 'GO' game includes a board with a 19x19 grid pattern, black and white stones, and two bowls to hold the stones. There are two players, one for each stone color. The player using black stones starts the game and then each opponent alternately places one stone on the grid intersections, trying to extend control over territory on the board. The game ends up with the winner who takes control over all possible territory. Players are not allowed to move existing stones to new locations, but a group of stones can be captured and removed from the board if all free connections are blocked by the opponent.

(a) (b)

Fig. 5 Demonstrating a compact HMPD prototype: (a) Front view of the prototype; (b) Image viewed through the exit pupil of the HMPD

3.2. Simulation outline

The simulation outline is illustrated in Fig. 6. Through an HMPD, a computer-generated 3D 'GO' board is projected onto a tabletop retro-reflective screen. The local player, wearing the HMPD, perceives the virtual board as if it was a real object sitting on the tabletop. He/she places real stone pieces on the virtual board. A head tracking system is used to maintain the correct registration of the real and virtual elements. A vision-based tracking setup detects the locations of the physical stones placed on or removed from the virtual board and communicates this information, via the collaborative server, to a remote player.

To simplify the simulation, we assume the remote player uses a PC-based game interface in which all game components are visualized on a PC monitor and stone manipulation is achieved via a standard mouse. When the player adds a piece to his/her board, the location of the piece is sent to the HMPD player via the collaborative server and a corresponding computer-generated piece is projected onto the HMPD user's virtual board.

Therefore, the HMPD player perceives the virtual board, his/her own real pieces, which correctly occlude the virtual board, and the virtual pieces placed by the remote player in a seamless augmented environment. Eventually, the remote player can have a similar 3D HMPD interface if visualization facilities are available at his/her site, or a video stream of the 3D visualization is transported across a suitable network.

Major technical aspects of the implementation will be described in section 4, calibration issues will be discussed in section 5, and experimental results will be included in section 6.

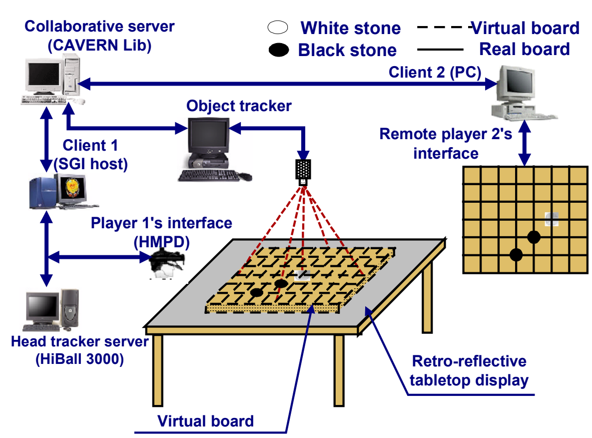


Fig. 6 Setup illustration of the testbed : playing augmented "GO" game with a remote opponent.

4. Testbed implementation

4.1. Framework for collaborative AR

We are developing a cross-platform application programming interface (API) specifically to support multiple-user remote collaboration in a 3D augmented environment [15]. The toolkit provides the base classes for the creation and visualization of a hierarchical tree of virtual scene objects, with related actions, and the classes for the rendering of the graphical contexts into multiple viewing windows. The API offers high-level network classes, supporting remote collaboration among arbitrarily many users. The library also provides base classes essential for the management of interface devices, such as head and object trackers, and data gloves. The library includes the associated functions for the display calibration of stereo viewing transformations and the registration of the virtual environment with its physical counterpart. Within the context of multiple-user collaboration, we will also expand the library to include base classes for the management of streamed video. Subjects of acquisition might include the AR visualization itself, the facial expressions of the collaborators, and the surrounding physical environments.

The 'GO' game embodies the following representative features of the API:

- Assembly and visualization of the scenegraph: The virtual 'GO' board and virtual stones are geometrically simple to model, but atrivial to position in the physical space. The number and location of virtual stones are event-dependent, with new stones being added or removed in arbitrary sequences.
- Collaboration via client/server architecture: Built atop the CAVERNSoft library [16], a separate freeware toolkit for tele-immersive collaboration, an independent server application routes the board coordinate data, for added and removed stones, between game clients.
- Control of user interaction: Three aspects of interaction are involved in this simulation. 1) A HiBall 3000 optical tracker is used to update HMPD user's view of the virtual components, therefore maintaining correct registration of virtual and real elements; 2) A vision-based object tracker locates, with respect to the virtual board, the physical stones that are added or removed by HMPD player; 3) A 2D, mouse-only interface drives interaction for the remote player. Currently, the remote manipulation of virtual or real stones placed by a remote opponent is not supported. Therefore, the capture action is made by the local participant himself. More interactive capabilities will be added in the future.
- Augmentation calibration: The intrinsic and extrinsic parameters of the viewing optics, as well as viewing transformations, are carefully calibrated to achieve dynamic registration of the virtual and real components. In the API, we implement the associated functions to establish a more accurate computational model for graphics generation.

The focus of the rest of the paper will be the implementation of the vision-based object tracker and the discussion of augmentation calibration.

4.2. Vision-based object tracker

Like many other AR applications, to achieve dynamic registration with their virtual counterparts or to support remote collaboration, an object tracker is required to detect the position or orientation of the physical objects of interest. In the 'GO' game case, the registration is maintained by the head-tracker, and the 2D locations of the physical stones, placed by the HMPD player, must be detected to update the view of the remote player.

Illustrated in Fig. 7(a), a vision-based object tracker determines the stone location relative to the virtual board. Infrared lamps are installed around the tracking camera to achieve coaxial illumination. Low-level coaxial illumination makes the retro-reflective background extremely bright. In a dimmed room-light condition, however, the image of the physical stones, either black or white, looks much darker than that of the retro-reflective background due to the diffusing property of the physical

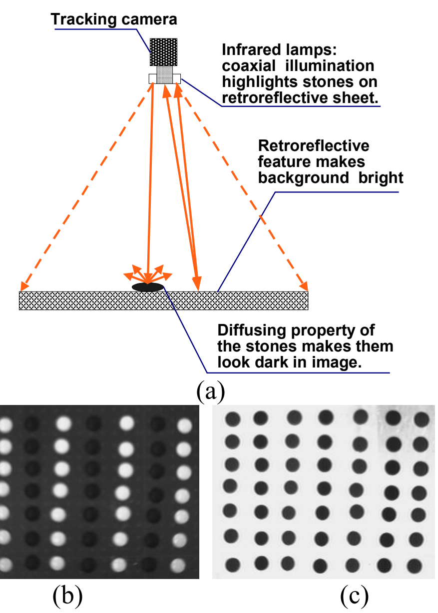


Fig. 7 A vision-based object tracker to detect physical stone locations: (a) Setup illustration: coaxial illumination highlights diffusing stones from retro-reflective background; (b) Image of black and white stones on retro-reflective background in normal lighting condition; (c) Image of black and white stones on retro-reflective background in dimmed room light with co-axial illuminators.

stones (Fig. 7(b)-(c)). Therefore, the coaxial infrared illumination distinguishes stones from their bright background. This feature allows us to robustly but efficiently locate the centroids of the real stones by using a simple algorithm.

Off-line calibration is made to determine the scale factors for mapping from image coordinates to virtual 'GO' board coordinates. The scale factors, measured on the image coordinates, include the pixel distances of two adjacent grid intersections in both horizontal and vertical directions, as well as the relative orientations of the virtual board on the sensor. The calibration is done by: 1) registering three physical stones with three corners of the virtual board; 2) detecting the centroids of the stones; 3) computing the scale factors through the image coordinates of the centroids.

To determine the location of each add-in or captured stone in the image coordinates, two frames of images are needed: the reference frame and the current frame. The reference frame, taken with no stones placed on the board, is subtracted from the current one to offset background noise, which makes the tracker usable in wide range of lighting conditions. The absolute intensity of the resultant image is a black background with discrete white spots. The spot centroids are computed to represent the locations of the stones in the image coordinates. Then the scale factors are applied to map the pixel coordinates to the virtual 'GO' board coordinates. A flag is assigned to differentiate the action of adding a new stone from the action of capturing an existing stone. In the current implementation, the player is required to press a trigger button after he/she adds or removes a stone to ensure his/her hand is out of the camera's view, but in the future, a function will be implemented to recognize this intuitively.

This tracking method is accurate and robust enough for the 2D tracking purpose required in the 'GO' application scenario. It allows the user to interact with the environment naturally, without any special attachment or marker on his/her hand and on the tracked objects. While the current tracking method is implemented for the 'GO' game, we are interested in extending it to a generic wireless 3D tracking method by using multiple cameras.

5. Calibration issues

Registering a virtual object with its real counterpart accurately and comfortably has been challenging in AR applications in the sense that the size and depth of the virtual objects have to be rendered precisely relative to a physical reference. The challenges in the 'GO' game are to ensure the virtual board aligns with the physical retro-reflective tabletop, and to ensure the virtual board appears in a fixed position and size in the real world space, when looking from arbitrary perspectives. Inaccurate or inconsistent depth/size perception can cause not only inaccurate registration, but also diplopic vision and eyestrain.

To ensure viewing comfort, since the current stereoscopic display technologies, including HMPDs, have a limited tolerance on viewing depth range, the average viewing depth and depth range of a specific application, which are related to accommodation/convergence conditions in stereoscopic displays, have to be determined empirically, and the display hardware has to be adjusted to comply with the specified conditions.

To precisely depict spatial relations between real and virtual objects, careful calibrations must be carried out to obtain accurate viewing orientation and projection transformations that are used to generate the virtual image pair [17]. Inaccurate representation of viewpoint position and viewing direction in the world space leads to errors of viewing orientation matrix, while inaccurate determination of display parameters causes errors in projection transformation matrix.

5.1. Accommodation/convergence considerations

Accommodation is a process of changing the focus of the eye for objects of varying distance, and convergence is a process where the eyes turn toward each other to aim the pupils directly at an object [18]. To avoid diplopic vision, in nature, humans have become accustomed to focusing and converging their eyes in a coordinated fashion: changing the distance from an observed object leads to not only changing the convergence, but also changing lens power. Stereoscopic displays, however, require the eyes be accommodated on the image plane, while they converge on a point of interest whose apparent depth may be different from the image plane. Under this condition, research by Valyus found that most people can comfortably tolerate a change in convergence angle of up to $\pm 1.6^{\circ}$ [19, 20]. Exceeding this tolerance leads to excessive parallax, called an accommodation/convergence conflict.

Setting the near and far clipping plane distance of the viewing frustum to limits specified by the convergence tolerance can eliminate the possibility of excessive parallax. For example, by using the maximum change in convergence angle of $\pm 1.6^{\circ}$ degrees, and 65mm of IPD, Southward recommended a viewing distance of at least 2.3m for viewing distant objects stereoscopically, which provides a comfortable stereoscopic depth range from about 1.2m to infinity [20]. Decreasing the viewing distance can bring the comfortable viewing range closer to the viewer in arm-length applications, for example, the 'GO' game, but this implies the viewer can tolerate a much smaller comfortable viewing range.

In the 'GO' game, if assuming the eye position is 0.5m away from the virtual board and the size of the virtual board is 0.45m, then the depth range of the virtual board is from 0.5m to 0.7m. Therefore, we set up the viewing distance of the display as about 0.6m, and the comfortable depth range is from 0.48m to 0.8m. Near and far clipping planes are set up accordingly when constructing the viewing projection transformation.

5.2. Determining viewing orientation transformation

Viewing orientation parameters specify the position and orientation of eyepoints in the world space. These parameters include the tracker position/orientation in the world space, $T_{W \leftarrow T}$, the position/orientation of the moving sensor in the tracker space, $T_{T \leftarrow S}$, and the left/right pupil/eye position and orientation in the sensor space, $T_{S \leftarrow E}$. Therefore, the viewing orientation transformation is given by:

$$T_{view} = T_{W \leftarrow E} = T_{W \leftarrow T} T_{T \leftarrow S} T_{S \leftarrow E}$$

While the head-tracker measurement gives the position/orientation of the sensor in the tracker space, $T_{T \leftarrow S}$, it is necessary to calibrate the position/orientation of the tracker coordinates in world space, $T_{W \leftarrow T}$. Typically, the actual origin and orientation of the tracker coordinates are not accurately aligned with those specified by the manufacturer, and therefore it is not possible to measure them directly. Our head-tracker is the HiBall 3000 from 3rd Tech, which comes with a stylus. In a designated world coordinate system, when aligning the tip of the stylus with the origin, a selected point along the X axis, and a point along Y axis of the world coordinate system, we recorded the position measurements of the sensor, $\vec{W}$, $\vec{X}$, and $\vec{Y}$, respectively. The normalized vectors of the X, Y, and Z axes in the tracker coordinate system are given by

$$\vec{X'} = \frac{\vec{X} - \vec{W}}{\left\| \vec{X} - \vec{W} \right\|}, \quad \vec{Y'} = \frac{\vec{Y} - \vec{W}}{\left\| \vec{Y} - \vec{W} \right\|}, \quad \vec{Z'} = \frac{\vec{X'} \times \vec{Y'}}{\left\| \vec{X'} \times \vec{Y'} \right\|}$$

Therefore,

$$T_{T \leftarrow W} = \begin{bmatrix} \vec{X'} & \vec{Y'} & \vec{Z'} & W \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

In terms of $T_{S \leftarrow E}$, in most of the applications, it is taken for granted that the left/right optical axes are parallel and point straightforward when the head stays upright, and therefore the viewing axes of the left and right eyes are assumed to be parallel. Practically, however, the left and right viewing axes can not be guaranteed parallel. Plus, in some HMD designs, the left/right optical axes are diverged or converged to achieve larger total FOV. Ignoring the non-parallelism can cause considerable perception errors. To more accurately model the $T_{S \leftarrow E}$, a calibration method, which is similar to the camera calibration methodology extensively used in the computer vision domain [21] but uses the special imaging property of the HMPD, is proposed and illustrated in Fig. 8. The HMPD is mounted at a fixed position/orientation and the head-tracker measurement is recorded $(\vec{S}, \vec{\Theta_S})$. An N-grid pattern is drawn on the retro-reflective screen, and the grid positions relative to the tracker coordinates are measured with the HiBall stylus, $(\vec{P_i})_T (i = 0, N-1)$. Observing at the exit pupil of the HMPD (or placing a camera in the eye location), the subject aligns a virtual cross on the display with each of the grid intersections and records the corresponding pixel positions of the cross on the image $(\vec{P_i})_I (i = 0, N-1)$. Using the least-square fitting method, we are able to compute the intrinsic and extrinsic parameters of the projective optics of the HMPD, such as the focal length, FOV, radial distortion coefficients, optical axes orientation, pupil/eye position, and display offsets. Applying this method to the left and right arms of the HMPD separately, their intrinsic and extrinsic parameters are listed in Table I.

Knowing the position/orientation of the calibration pattern as well as the motion sensor in the tracker space, we further compute the transformation $T_{S \leftarrow E}$ to give the pupil/eye position/orientation in the sensor space. The transformations for the left and right arms are:

$$[T_{S \leftarrow E}]_{Left} = \begin{bmatrix} 0.9996 & 0.0044 & 0.0264 & -36.1664 \\ -0.0049 & 0.99981 & 0.0208 & -147.2725 \\ -0.0263 & -0.0209 & 0.9994 & -21.4114 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

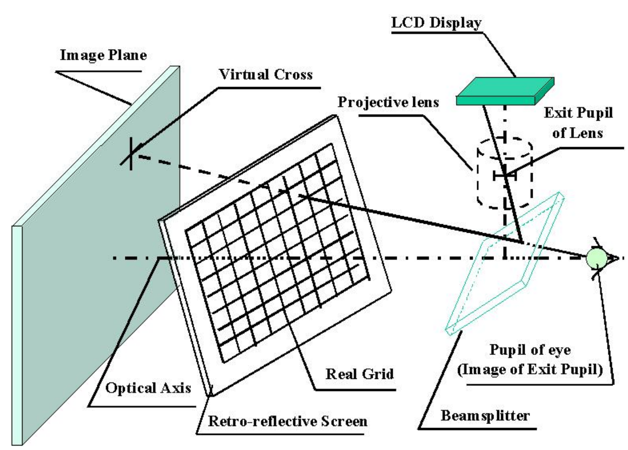


Fig. 8 Illustration of the HMPD calibration method

$$[T_{S\leftarrow E}]_{Right} = \begin{bmatrix} 0.9999 & 0.0102 & 0.0069 & 32.4025 \\ -0.0097 & 0.9981 & -0.0616 & -135.3587 \\ -0.0075 & 0.0615 & 0.9981 & -7.7901 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Table I Intrinsic/extrinsic parameters of the HMPD

Parameters	Left	Right
Focal length (mm)	35.1473	35.09793
FOV (degrees)	41.17(H) 31.76(V)	41.22 (H) 31.81(V)
Distortion coefficient (1/mm^2)	8.134e-005	1.310e-004
Pupil position relative to calibration pattern (mm)	X=-178.716 Y=-171.653 Z=773.7375	X=-214.9796 Y=-105.5128 Z=688.77537
Orientation of the optical axis (degree)	Yaw=143.35 Pitch=-4.73 Roll=90.37	Yaw=142.299 Pitch=0.0061 Roll=90.1573
Display offsets(pixel)	H=77 V=-55	H=14 V=21

Therefore, in our implementation of the viewing orientation transformation, we not only take into account the calibration of the world-tracker transformation, but also implement two separate eye-to-sensor transformations for the left and right viewpoints, specifically compensating for the pupil offsets relative to the sensor coordinates and the variation of the left/right optical axis orientations.

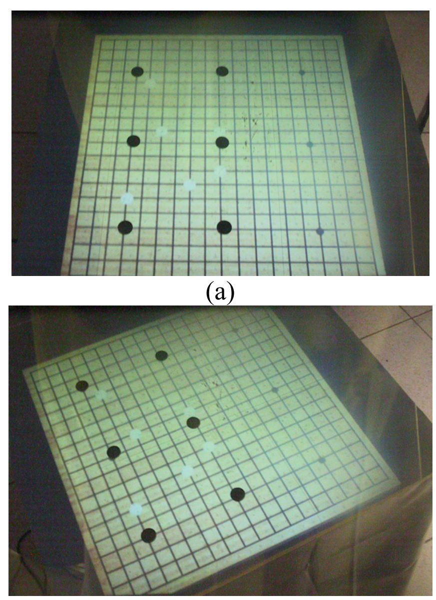
This technique assumes that the optical pupils will always match with the eyes of all users (interpupillary distance variation is considered separately) and does not consider the possibility that the eye position relative to the optical pupils may vary with different users or change during the application due to helmet slippage.

5.3. Determining viewing projection transformation

The viewing projection transformation specifies how a 3D virtual object in the world space is projected onto a 2D viewport. Most of the graphics packages assume that the viewing direction is normal to the viewing plane [17]. To be compatible with this assumption, we assume the optical axis is normal to its associated image plane, and its orientation relative to sensor space is taken into account in the viewing orientation transformation. Therefore, the viewing projection parameters include the field of view (FOV), view frustum and frustum asymmetry, interpupillary distance (IPD), and optical distortion. These relevant parameters are obtained through the calibration process described in the last section. Optical distortion is ignorable in our HMPD prototype due to a very good correction in the projection lens design. In our implementation, the view plane, and the far and near clipping planes are set up to the image distance of the display, and the farthest and nearest depth limits specified by convergence tolerance, respectively. We implement two separate series of viewing volume asymmetries for left and right eyes, respectively. For each eye, we consider both the horizontal and vertical asymmetries, which are set to be the horizontal and vertical offsets of the intersection of optical axis with the image plane from the image plane center. We also ensure the computational baseline of the left/right virtual cameras match with the user's IPD as well as the baseline of the binocular display.

5.4. Evaluation

By implementing the calibration methods in our software, we achieved much more accurate rendering than the case without those calibration procedures. Without making any other adjustment, the perceived virtual 'GO' board is fairly well aligned with real stones placed on the top of the retro-reflective screen at changing perspectives. Fig. 9 shows the registration at two different perspectives. Subtle discrepancy of the perceived depth/size from the designated values still exists and slight variation of the perceived location also exists when observing at changing perspectives. The current calibration method is time-consuming and can only be done off-line. Currently, we do not consider the possibility of mis-positioning of the helmet for different users. We also do not consider the possibility of perception artifacts caused by the presence of retro-reflective screen. Furthermore, the registration is mainly achieved on a 2D plane, not accurate enough for



(a)

(b)

Fig. 9 Black real stones properly register with the virtual elements and occlude the virtual board at different perspectives in HMPD: (a) The front perspective; (b) Side perspective.

more demanding applications in 3D. More accurate calibration methods and computational models are required to fully compensate for all possible error sources. Further investigation on generic calibration methods is under way.

6. Experimental results

Figure 10 (a) shows the simulation setup, and figures 10 (b) through (e) show the virtual and direct views of both players. Figure (b) depicts the virtual components seen by the HMPD player, including a virtual board with the white stones placed by his/her remote player. Figure (c) shows the HMPD player's direct view, with only his physical stones scattered on the screen. Figure (d) shows the augmented view perceived through the HMPD: the virtual board, white virtual stones, black real stones, and miscellaneous elements of the physical environment are seamlessly integrated, with the black stones naturally occluding the occupied grids. Figure (e) shows the view of the remote player. As a testbed, the 'GO' game demonstrates the augmentation and registration of real and virtual objects, and natural occlusion of virtual objects by real counterparts, as well as interaction and remote collaboration with a remote participant using HMPD technology.

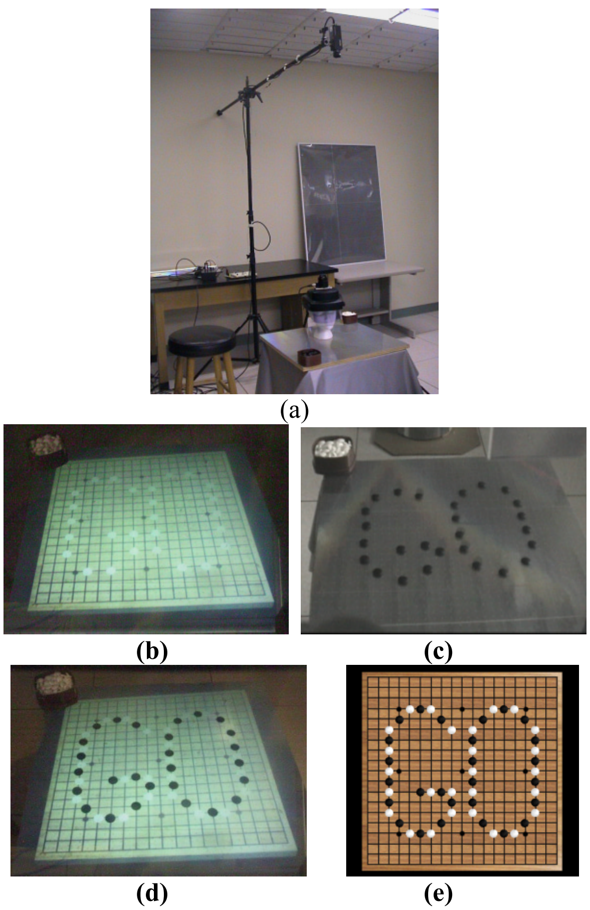


Fig. 10 Playing "GO" game with a remote opponent: (a) A close-up of the setup; (b) HMPD player's virtual view; (c) HMPD player's direct real view; (d) HMPD player's augmented view; (e) Remote player's PC-based interface.

7. Conclusion

The head-mounted projective display (HMPD) has been proposed as an alternative solution to optical see-through devices. Its main advantages include the capabilities of: 1) achieving larger FOV and easier correction of optical distortion than conventional eyepiece-based optical STHMDs; 2) allowing correct occlusion of virtual objects in augmented environments; 3) projecting undistorted images on curved surfaces at arbitrary position; and 4) creating independent viewpoints without crosstalk in multi-user environments. In this paper, we first reviewed the featured capabilities of the HMPD technology and the recent development in both display implementations and applications. Then we presented the implementation of our first testbed, namely "playing 'GO' game with a remote opponent in a 3D augmented environment". Moreover, through the testbed, important calibration issues of the HMPD, such as accommodation/ convergence considerations, and size/depth perception, were studied. Calibration methods and results were discussed in detail, which are applicable to other applications. Finally, experimental results of the testbed implementation were demonstrated. The 'GO' testbed demonstrated the capabilities of virtual-real augmentation and registration, natural occlusion of virtual objects by real, interaction with an augmented environment, as well as networking collaboration. It also embodied part of our long-term objective to develop a framework to support remote collaboration in 3D augmented environments. In future work, more accurate calibration methods will be investigated and comprehensive evaluation on the perception performance will be carried out.

Acknowledgements

We would like to specially acknowledge our collaborators Dr. Frank Biocca of the M.I.N.D Lab at Michigan State University, and Dr. Ken Smith from 3M. This paper is based on work supported by National Science Foundation Grant IIS 00-83037 ITR and Beckman Institute Equipment Allocations.

References

[1] Ivan Sutherland. "The ultimate display". *Proceedings of IFIP 65*, Vol. 2, 1965, pp.506-508.

[2] D. Buxton, G W. Fitzmaurice. “HMDs, caves and chameleon: a human-centric analysis of interaction in virtual space”. *Computer Graphics (ACM)*, Vol. 32, No. 4, Nov. (1998), 69-74.

[3] Rolland, J.P., and H. Fuchs, "Optical versus video see-through head-mounted displays in medical visualization," *Presence: Teleoperators and Virtual Environments (MIT Press),* 9(3), (2000), 287-309.

[4] J. Fergason. “Optical system for head mounted display using retro-reflector and method of displaying an image”, U.S. patent 5,621,572. April 15, 1997.

[5] J. Parsons, and J. P. Rolland, “A non-intrusive display technique for providing real-time data within a surgeons critical area of interest,” *Proceedings of Medicine Meets Virtual Reality98*, 1998, 246-251.

[6] Hong Hua, A. Girardot, David Poizat, and J. P. Rolland. “Head-mounted projective display: design and engineering study”, *Oral presentation at OSA 1999 Annual Meeing/XII-XL*, San Jose, Sep.23-27, 1999 (Abstract).

[7] Hong Hua, A. Girardot, Chunyu Gao, and J. P. Rolland. “Engineering of head-mounted projective displays”. *Applied Optics*, 39 (22), 2000, 3814-3824.

[8] Hong Hua, Chunyu Gao, Frank Biocca, and Jannick P. Rolland, "An Ultra-light and Compact Design and Implementation of Head-Mounted Projective Displays", *Proceedings of IEEE-VR 2001*, p. 175-182, March 12-17, 2001, Yokohama, Japan.

[9] J. P. Rolland, J. Parsons, D. Poizat, and D. Hancock, “Conformal optics for 3D visualization,” *SPIE-Int. Soc. Opt. Eng. Proceedings of SPIE-the International Society for Optical Engineering*, vol.3482, 1998, pp.760-4.

[10] Naoki Kawakami, Masahiko Inami, Dairoku Sekiguchi, Yasuyuki Yangagida, Taro Maeda, and Susumu Tachi, “Object-oriented displays: a new type of display systems—from immersive display to object-oriented displays”, *IEEE SMC'99 Conference Proceedings*, 1999 IEEE International Conference on Systems, Man, and Cybernetics, Vol.5, 1999, pp.1066-9 vol.5. Piscataway, NJ, USA.

[11] Masahiko Inami, Naoki Kawakami, Dairoki Sekiguchi, Yasuyuki Yanagida, Taro Maeda, and Susumu Tachi, “Visuo-haptic display using head-mounted projector”, *Proceedings IEEE Virtual Reality 2000*, IEEE Comput. Soc. 2000, pp.233-40. Los Alamitos, CA, USA.

[12] Ryugo Kijima and Takeo Ojika, “Transition between virtual environment and workstation environment with projective head-mounted display”, *Proceedings of IEEE 1997 Virtual Reality Annual International Symposium,* IEEE Comput. Soc. Press. 1997, pp.130-7. Los Alamitos, CA, USA.

[13] Jannick P. Rolland, Hong Hua, and Frank Biocca, “Compact Lens-Assembly for Wearable Displays, Projection Systems, and Cameras”, *Patent Application. United States: University of Central Florida.*

[14] Hong Hua et. al., “A Head-Mounted Projective Display and its Applications in Interactive Augmented Environments”, in *SIGGRAPH 2001 Conference Abstracts and Applications--Sketches & Applications*, Aug. 12-17th, 2001, Los Angeles.

[15] Hong Hua, et. al, “Using a head-mounted projective display in interactive augmented environments”, in *Proceedings of IEEE International Symposium on Augmented Reality 2001*, Octomber 29-30, 2001, New York, NY.

[16] http://www.openchannelsoftware.org/projects /CAVERNsoft_G2/

[17] J. D. Foley, A. van Dam, S. K. Feiner, and J. F. Hughes, Computer Graphics: Principles and Practice, Addision-Wesley, Reading, MA, 1990.

[18] I. P. Howard, B. J. Rogers. Binocular vision and stereopsis. Oxford psychological series n.29, New York: Oxford University Press, 1995.

[19] N. A. Valyus, Stereoscopy, Focal Press, London, 1996.

[20] David A. Southard, Transformations for stereoscopic visual simulation, *Computer & Graphics* Vol.16, No.4, pp.401-410, 1992.

[21] Roger Y. Tsai. “A versatile Camera Calibration Technique for High-Accuracy 3D Machine Vision Metrology Using Off-the-Shelf TV Cameras and Lenses”, *IEEE Journal of Robotics and Automation*, Vol. RA-3, No. 4, August 1987, pages 323-344.