

# A Compact Optical See-through Head-Worn Display with Occlusion Support

Ozan Cakmakci, Yonggang Ha and Jannick P. Rolland  
*College of Optics/CREOL*  
University of Central Florida  
{ozan,ha,jannick}@odalab.ucf.edu

## Abstract

*We are proposing a novel optical see-through head-worn display that is capable of mutual occlusions. Mutual occlusion is an attribute of an augmented reality display where real objects can occlude virtual objects and virtual objects can occlude real objects. For a user to achieve the perception of indifference between the real and the virtual images superimposed on the real environment, mutual occlusion is a strongly desired attribute for certain applications. This paper presents a breakthrough in display hardware from a mobility (i.e. compactness), resolution, and a switching speed based criteria. Specifically, we focus on the research that is related to virtual objects being able to occlude real objects. The core of the system is a spatial light modulator (SLM) and polarization-based optics which allow us to block or pass certain parts of a scene which is viewed through the head-worn display. An objective lens images the scene onto the SLM and the modulated image is mapped back to the original scene via an eyepiece. We are combining computer generated imagery with the modulated version of the scene to form the final image a user would see.*

**Keywords:** display hardware, occlusion, augmented reality, optical system design, head mounted display, spatial light modulator

## 1. Introduction

Augmented reality can be defined as adding or subtracting information to the human senses. In general, any human sense can be augmented. In this paper, we are referring to the augmentation of the visual sense. One flavor of AR is based on the use of head-worn displays. One could classify head-worn displays into two classes: optical and video see-through[1]. There are many scientists who prefer video see through, mostly for technical reasons. For example,

using chroma keying it is easy to add information to the scene, and also it is relatively easy to block out certain parts of the scene. From a human usability point of view, there are several well-known issues with video see-through AR. Lag due to processing of the incoming video stream is a big issue in video see through AR. Our system is capable of providing the technical attributes that make video-see through displays attractive in an optical manner. Optical see-through displays are much faster than video see-through displays and they provide excellent resolution of the real scene given that the scene is not sampled by cameras, rather the human eye gets a direct view of the real world.

## 2. Related Work

For optical see-through displays, starting with Ivan Sutherland's original head-worn display, most optical designs[6], even today, will combine the computer generated imagery with the real world using a beam splitter. Sutherland concludes in his 1968 paper that "showing "opaque" objects with hidden lines removed is beyond [their] present capability"[7]. He adds that "the three-dimensional objects shown by [their] equipment are transparent "wire frame" line drawings". It is well established that any beamsplitter will reflect some percentage of the incoming light and transmit the rest. The percentage of transmission and reflection can be adjusted through coatings on the surfaces. However, there will always be some amount of light that is transmitted from the scene which is the root of the occlusion problem in optical see-through displays. This transmitted light implies that it is optically impossible to overlay opaque objects for a display using a combination of image source and a beamsplitter. We will illustrate the concept of occlusion graphically within the next section. Alternative mechanisms (to the

beam splitter + image source combination) will be necessary to achieve the continuum between transparency and opaqueness of virtual objects. To be concrete about this transparency and opaqueness continuum of virtual objects, a virtual object needs to be opaque when it is occluding and it needs to be transparent when it is occluded by an object in the real world scene. Current computer graphics techniques and hardware allow for "hidden line removal" or visible surface determination, however, the display of "opaque" objects with optical see-through remains a problem. Recently, Bimber presented a projector based technique to the occlusion problem [14]. The basic idea consists of projecting shadows to produce correct occlusion effects between the real and the virtual objects.

Ian Robinson disclosed the concept for a system where a transmission type device was utilized to block/pass certain parts of the scene [3]. Soon after, in 1999, an article from Eric W. Tatham [16] illustrated results from a transmissive light blocking arrangement with no imaging optics. An active mask was proposed to modulate the content of a scene and was combined with the display. Tatham further pointed to some of the potential benefits of using a Digital Micromirror Device (DMD) in place of the transmissive mask, yet no optical layout was proposed. We were unable to find further discussion of this idea from either of the authors in the form of prototypes within conference proceedings or journals. However Kiyokawa and Billinghurst in 2000 addressed the occlusion problem in their ELMO prototype, now in its fourth generation [2]. Vivid images of mutual occlusion were first demonstrated by the authors. The heart of their system is a transmissive spatial light modulator (SLM). ELMO-4 is based on a 320x240 transmissive LCD from Hunet. The LCD is reported to have a response time of 2ms. The ELMO-4 optical system contains 4 lenses, 2 prisms, and 3 mirrors per eye for the display component. The designers of the ELMO-4 prototype took special care to eliminate the viewpoint offset caused by the optics which justifies the current size for the approach taken by the authors. The other components in the ELMO-4 are additional cameras located around the optics to enable depth extraction of the scene, which will not be further addressed in this paper. Reflective SLM approach using a DMD has been pursued by Uchida, Sato and Inokuchi [4]. According to our understanding (original paper is in Japanese), in this prototype, there are three optical paths. One optical path is for imaging the scene onto the DMD, another one is for imaging the microdisplay onto the DMD and the last one is for projection and/or

eyepiece optics. This is a promising prototype due to the high contrast ratio of the DMD device. However, further work on the optical layout will be needed to combine the three separate optical paths into a single path and also to eliminate the potential viewpoint offset, if this approach is to be used in a head-worn display.

We are proposing in this paper a novel optical see-through head-worn display that is capable of mutual occlusions[15]. A key difference between our system and the previous systems is that our system is based on a reflective SLM combined with a single optical path, suitable for use in head-worn displays. Our system has a resolution of 1280x1024, a switching speed within microseconds, which is due to the faster switching in ferroelectric liquid crystal compared to nematic liquid crystals [10] (an order of magnitude faster than the transmissive masks), a 60Hz video rate, and importantly is optically compact (e.g., 2 lenses and a single x-cube prism per eye). The reason our design yields a compact form is that it does not require erecting optics, rather it uses polarization-based optics as shown in Section 4. A crude estimate of the weight of the proposed head-worn display yields less than 50 grams per eye. This estimate accounts for 17 grams for the X-cube prism, 5 grams for the lenses, and ~30 grams for the LCOS and the driver electronics. In the transmissive mask approach of ELMO-4, the LCD module was reported to have a contrast ratio of 1:100, whereas an F-LCOS (ferroelectric liquid crystal on silicon), which we propose to use as one implementation of the SLM panel, yields >1:200 contrast ratios [11]. Also, we expect a light throughput between 40-50% of what the scene would provide for the real world objects. For the virtual objects, we expect a 50-100% light throughput depending on whether the microdisplay used to display the images is polarized (e.g. LCD and F-LCOS panels, the later being a possibility to be used also as a display) or non-polarized (e.g. OLEDs). Based on 700-1000 fL brightness of current microdisplay technology, we can expect a 350fL-1000fL brightness on the display of virtual objects. Even the lowest end of this range is enough for displaying information in bright indoor environments. Finally, we shall show that the system can be designed distortion free while viewing the real scene.

### 3. Illustration of the Occlusion Concept

A basic idea is that, depending on the optical properties of a material, in general the complex



**Figure 1. Concept of occlusion illustrated with a virtual apple as would be seen through a beam splitter based head-worn display**



**Figure 2. Concept of occlusion illustrated with a virtual apple as would be seen through an ideal occlusion based display.**

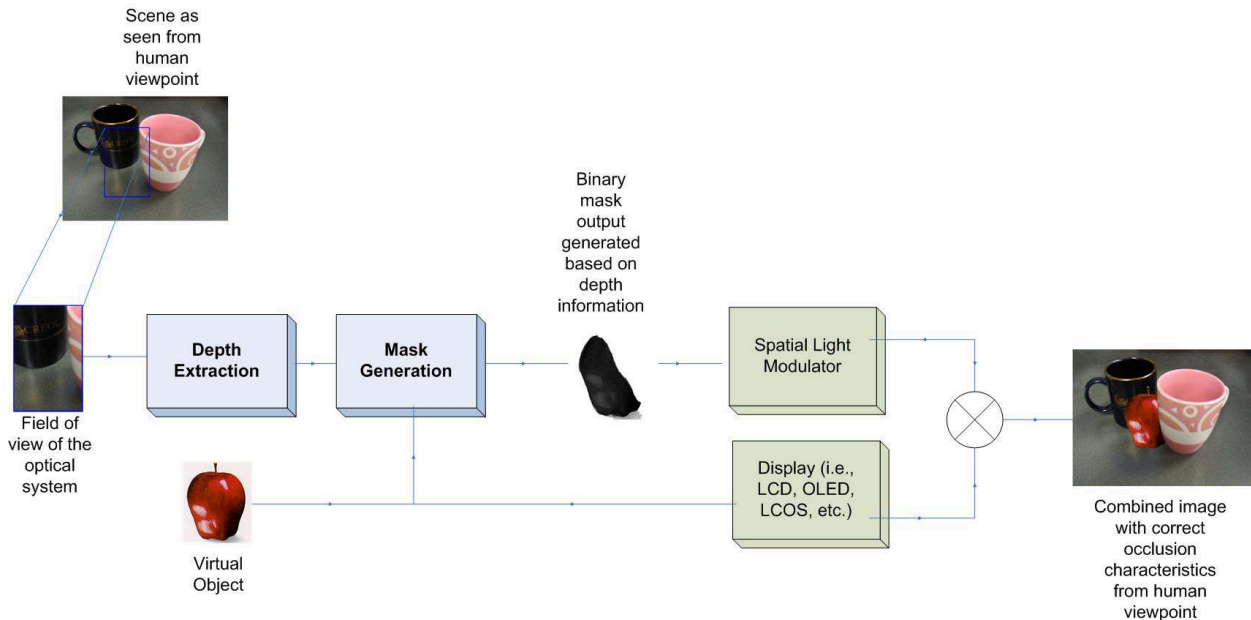
refractive index  $\eta = n + i \cdot \kappa$  of an object, light is transmitted through a material, reflected, or it is absorbed.  $n$  is the real part of the refractive index and  $\kappa$  is the complex part of the refractive index which relates to the absorption coefficient  $\alpha$  [ $\text{cm}^{-1}$ ]. Occlusion is a result of light propagation and attenuation through different materials within a scene.

Occlusion is a concept very familiar to us from daily life. For example, a tall person sitting in a theatre would block the light from the theatre stage for the people immediately in the row behind since light in the visible spectrum is not transmitted through people. We can say that the tall person is occluding the field of view of the people sitting in the immediate row behind.

We present a visual example of occlusion in the context of augmented reality. Fig. 1 illustrates a case

where a virtual apple would be overlaid on a real scene with a conventional head-worn display. We can observe that the apple is semi-transparent as a result of the transmission through the beamsplitter. Fig. 2 illustrates a virtual apple overlaid on a real scene through an occlusion supporting head-worn display. We can observe that the apple occludes the cup behind it and is occluded by the cup in front of it, consistent with our daily experience of objects mutually occluding each other.

Based on this ideal occlusion illustration shown in Fig. 2, we outline the issues to be addressed as follows and illustrate the steps that such a system would take in Fig. 3:



**Figure 3. Block Diagram of an Ideal Occlusion System**

- 1) An optical system that can image a desired amount of the field of view, based on the application needs.
- 2) A depth sensing mechanism.
- 3) An algorithm to create a binary mask based on depth information.
- 4) An optical system that modulates the scene based on the mask.
- 5) An optical system that combines the modulated scene with the image display hardware.

We envision the application scope of our head-worn display to include mobile outdoor scenarios. Mobility implies that the system should be compact and light weight for maximum comfort while wearing the display.

#### 4. First Order Layout for the Optically Compact Display system and the occlusion mechanism

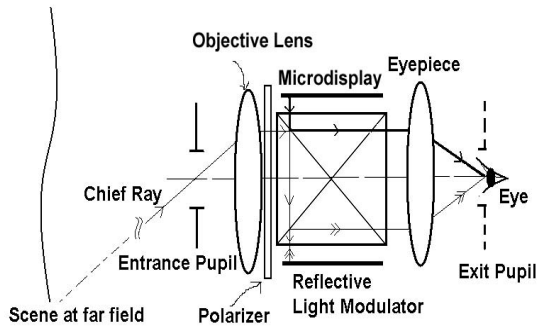


Figure 4. First order optical layout.

As depicted graphically in Figure 4, our system consists of an objective lens, a polarizer, an x-cube prism, a reflective SLM (e.g. F-LCOS, DMD), a microdisplay as an image source, and an eyepiece.

The objective lens collects the light from the scene, in a telecentric manner, and delivers it to the SLM for modulating the content of the scene (pass or block light from the scene). A lens is *telecentric* in a given space if the chief rays (i.e., rays in the field of view that pass through the center of the pupil – see Fig 4) are parallel to one another in that space. Most commonly, they are also parallel to the lens axis and perpendicular to the object and / or image planes that are perpendicular to the axis [8]. We will refer to the image on the SLM as an intermediary image since this image lies between the object and image spaces. In optics, the space an object lies in is defined to be the object space and an analogous definition exists for the image space. The SLM for this application can be modeled as a perfect flat mirror with a quarter wave-plate (i.e. an optical device that rotates polarization by

90 degrees in the case of double pass). The plate or an equivalent polarization rotator is required to properly handle polarization throughout the system in order to maximize light throughput and suppress ghost images as further discussed later in this Section.

In our system, telecentricity in the intermediary space is required for three reasons:

- 1) From a first order layout perspective, the system requires a 90 degrees angle of incidence for the chief rays in order for the SLM to operate efficiently. The optical light path is shown in Fig. 4. A deviation from the 90 degree requirement will cause vignetting (i.e. loss of light). Vignetting occurs when an image-forming bundle is truncated by two or more physical structures in different planes [8].

- 2) Since our design is based on an x-cube polarizing combiner, we do not want to come in to the x-cube with an angle. This is analogous to a converging beam of light passing through a plane parallel plate at an angle which will result in severe optical aberrations.

- 3) For a given chief ray angle, imaging of various object planes or defocus of the SLM will not change the image height which will facilitate performing the occlusion for a given region of the scene around the chief ray (defocus will only blur the image in a telecentric system).

Once the scene is modulated according to the application needs, the eyepiece will take the modulated light output and map it back to the original scene (the exact meaning of “mapping” is explained below). Even though the choice of delivering a collimated image to the eye is an option and not a requirement for the proposed system, in this design, we will discuss the case of a collimated image in visual space. This will be a justified choice for virtual objects displayed in the far field. To get a collimated image from the output of the SLM, we need to place the front surface of the eyepiece a front focal length away from the SLM. The height of the intermediary image is fixed by the height of the SLM, therefore if the focal length of the eyepiece matches the focal length of the objective, the field of view for both the objective and the eyepiece will be exactly the same. This is what we mean by mapping the modulated image back on to the scene. Matching the field of view of the objective with the field of the eyepiece will ensure 1:1 magnification.

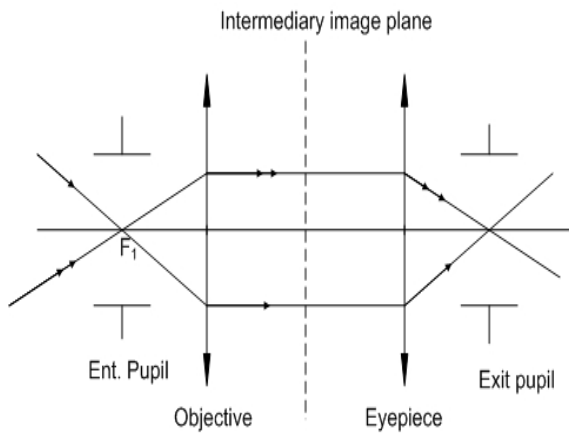
The human eye is conjugated to the entrance pupil which will cause a viewpoint offset on the order of 50mm. A simple stereo depth resolution model estimates that over  $\sim 3.3m$ , this offset will not be perceived by the eye. This estimate uses a stereo depth resolution value of  $1/2$  arcminute for the human eye[8]. When we consider the LCOS pixel size setting

the acuity limit, the result shows that after ~6m this offset will not be perceived by the eye.

Fig.5. shows an unfolded optical layout of the system shown in Fig. 4. It is typical to unfold reflective system. Unfolding is achieved by replacing a mirror having a particular optical power with a lens of the same power. The unfolded system is easier to analyze and think about compared to the system shown in Fig. 4.

The entrance pupil must be placed at the focal point of the objective to force the chief rays in the intermediary image space to be parallel to the optical axis, in order to make the system telecentric in that space. It is well known that whenever the chief rays cross the optical axis, we get a pupil in that space. Therefore, after the telecentric intermediary space, the chief ray will focus down at the focal point of the eyepiece. This will form the exit pupil of the system. Ideally, to accommodate people wearing eyeglasses, the exit pupil should be located around 25mm out from the last surface.

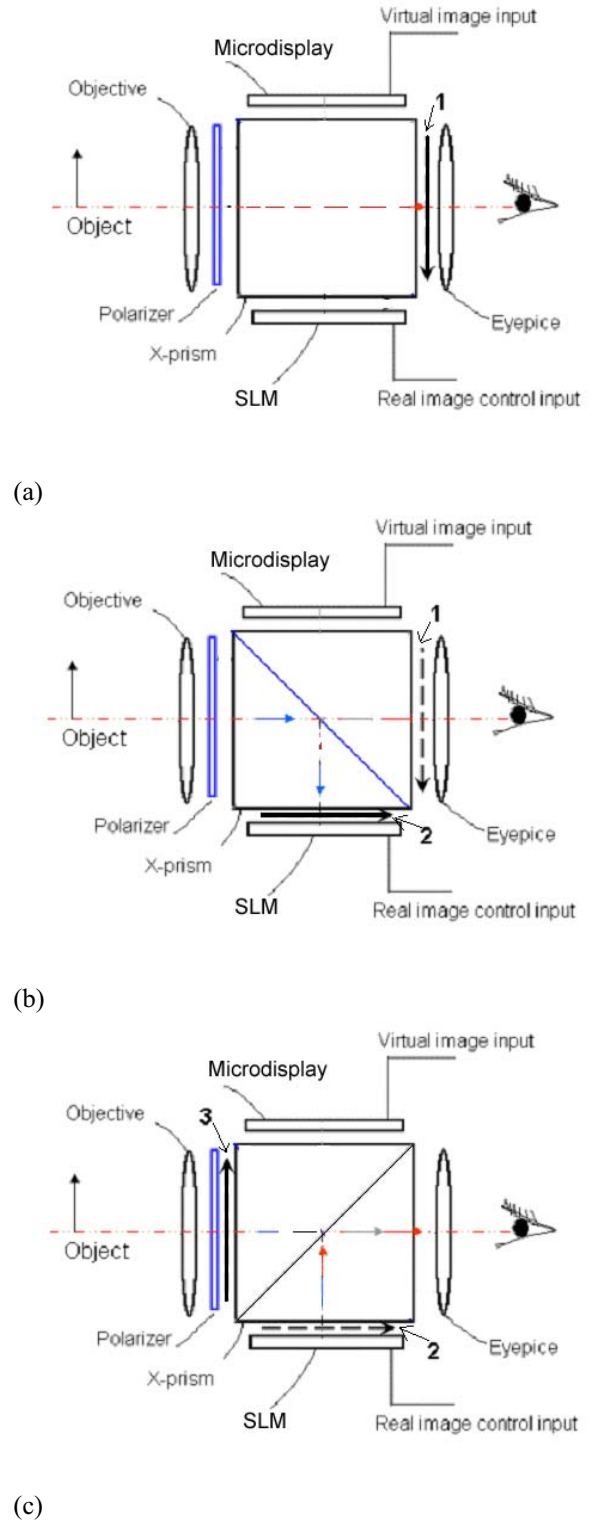
An exit pupil diameter of 9-12mm is desirable, within which the pupil of the eye (i.e., 2-4mm in diameter) will be located and naturally move within  $\pm 20$  degrees [13]. The larger the pupil of the instrument compared to the eye pupil, the larger the eye motions allowed. However this occurs at the expense of compactness. We propose a system with a 9-mm pupil to allow the  $\pm 21$  degrees natural eye motion together with maximum compactness, without comprising field of view.



**Figure 5. Unfolded light path with the chief rays raytraced through the head-worn optics.**

Next, we verify that the final image will have the desired upright orientation with respect to the eye. In order to verify orientation, we need to consider the

system shown in Fig. 4 to take into account the effect of reflection from a mirror on orientation (e.g., can't



**Figure 6. Verification of upright image orientation.**

use the unfolded system in this case). The diagram pertinent to verifying image orientation is shown in Fig 6a-c. The object is indicated with an upright arrow and it is assumed to have an initial upright orientation. The object is first imaged through the objective lens and has an inverted orientation as indicated in orientation at step “1” with a solid black line shown in Fig. 6a. Due to the polarizer, right after the lens, the light will be s-polarized, therefore, it will hit the s-reflect coating in the x-cube prism. The orientation upon reflection is shown in step “2” represented in Fig.6b as a solid black line close to the SLM. The SLM will reflect the image and change the polarization, assuming the pixel is “turned on”. Caused by this change of polarization, the light will now be p-polarized and therefore hit the p- reflect coating on the x-cube and will be directed towards the eye as shown in Fig. 6c. The orientation after the p-reflect mirror is shown in step “3” of Fig. 6c, the final step in the analysis. We can clearly verify that the final image will have an upright orientation.

## 5. Prototype Specification

We created a table of specifications for a prototype design implementation. The specifications for the objective and the eyepiece, which will be shown to be the same element by design, will be provided in Section 6.

We first designed the first order optical layout of the objective lens. The goal of this lens is to image a specified field of view onto the SLM for selectively turning on/off pixels in the real world scene. For our design goal, we are setting the field of view (FOV) of the objective lens to be 40 degrees (full field). Knowing the FOV and the size of the SLM, we find the focal length of the lens to be 30.7 mm. In an analogous manner, we find the horizontal and the vertical FOV to be + - 15.81 degrees and + - 12.77 degrees, respectively. Based on this scheme, the optics will not require erection of the image and can be very compact. Note that the size of the SLM and the microdisplay is about 1”, taking the reflections we need into account (45 degrees), the x-cube turns out to be a cube with a size 1” cubed. Finally, in this first feasibility investigation, we are setting the F/# (defined as the ratio between  $f$ , the focal length, and  $D$ , the diameter of the lens) to be 3.75 as shown in table 1. The F-LCOS, when used as the SLM, can support optical designs down to  $f/1.8$ , which will provide even brighter images.

The pixel period is on the order of 30 microns for the F-LCOS, therefore, the maximum spatial frequency will be 36.27 cycles/mm (i.e.  $1000/(2*13.62)$

cycles/mm). With this modulator, any image information with a higher spatial frequency (i.e. finer detail) will not be seen. The standard value of the MTF at the maximum spatial frequency is 20%, which we set for our design as shown in table 1.

## 6. Lens Design and User-Centric Assessment

For the optical design, we will use the same optics for the objective and the eyepiece. The advantage of using the same optics is to fully cancel out the distortion for the real scene, given that the eyepiece distortion will be of opposite sign to that of the objective by symmetry. Regarding the virtual image, the distortion will be that of the eyepiece, therefore we shall still limit that value to less than 10% and propose to compensate such distortion as needed for the application either in software or in hardware.

| Basic System Parameters          |  |
|----------------------------------|--|
| Object Distance                  | Infinity   |
| Eye relief (last surface to eye) | > 23 mm (eyeglasses compatible)<br>25mm was achieved in our design |
| F/#                              | < 3.5<br>3.41 was achieved   |
| Wavelength band                  | Visible (486-656 nm)   |
| Full Field of View               | 40 degrees desired   |
| MTF max spatial freq.            | 37 cyc/mm  |
| MTF at max spatial freq.         | 20%  |
| Distortion                       | < 10% (for the virtual scene)<br>0% for the real scene             |
| <b>Sensor type</b>               | F-LCOS modulator   |
| Full diagonal                    | 0.88” or 22.4 mm   |
| Number of Pixels                 | 1280x1024 pixels   |
| Horizontal size                  | 17.43 mm   |
| Vertical size                    | 13.95 mm   |
| Pixel size                       | 13.62 $\mu$ m x 13.62 $\mu$ m                                      |
| <b>Display type</b>              | LCD Display  |
| Diagonal size                    | 0.9”/23mm  |
| Number of pixels                 | 1024x768 pixels  |
| Horizontal size                  | 18.4 mm  |
| Vertical size                    | 13.8 mm  |
| Pixel size                       | 18 x 18 $\mu$ m  |

Table 1. Basic System Specifications.

To summarize the lens design goals, we want to obtain a physically compact relatively wide angle (40 degrees) telecentric objective that is separated from the F-LCOS by 25.4mm to allow the insertion of the x-cube. The telecentricity requirement imposed that the entrance pupil be outside of the system and this requirement inspired us to use a rear landscape lens.

An introduction to the landscape lens can be found in [5]. Importantly the landscape lens does not typically work in telecentric mode. So we expected upon optimization to have the lens evolve to a best form lens which has a shape close to a plano-convex singlet. However, the optical system also required to work with a wide spectrum and with a fairly large field of view. Therefore, we added a second element to help with aberration correction. The first element was kept as a conventional optical element with spherical surfaces and the second element was made into a diffractive optical element (DOE) with a single diffractive surface to balance chromatic aberrations.

In this way, both compactness and good image quality are achieved. Fig. 7 shows the layout of the two element system. The chief rays are shown to be parallel to the optical axis which achieves the telecentricity requirement.

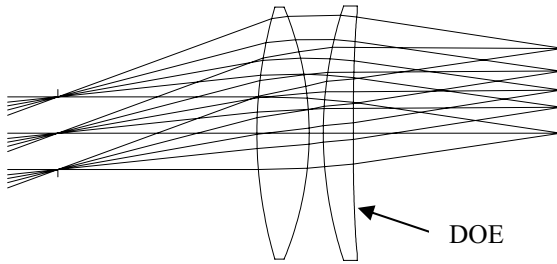


Figure 7. Layout of the optimized objective lens.

Modulation transfer function (MTF) is the magnitude response of the imaging system to sinusoids of different spatial frequencies[8]. This response is defined as a measure of modulation depth:

$$M = \frac{A_{\max} - A_{\min}}{A_{\max} + A_{\min}}$$

A is the value of the waveform that describes the object or the image. The MTF physically represents the contrast of the various spatial frequency objects after passing through the optics.

Fig. 8 shows the MTF of the objective lens for a 2mm pupil. The MTF holds relatively well across the field of view. We can observe that at the maximum spatial frequency, the modulation transfer function is

greater than 10% across the field of view. More optimization with weighting of the various FOVs can further help bringing the curves together and above a preferred 20% modulation at 37 cycles/mm.

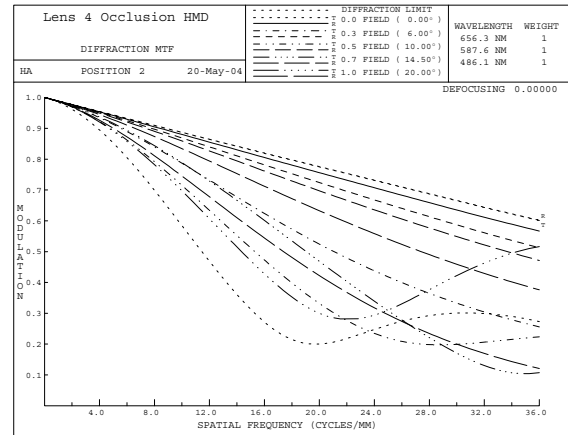


Figure 8. MTF of the two element system with a 2mm pupil

Fig. 9 shows the distortion grid of the two element system. At the edge of the field, the distortion is less than 8%. This is the distortion of the virtual scene. Prewarping the image on the microdisplay (e.g. LCD), we can correct for this distortion. As previously discussed, the real scene is not distorted due to the symmetry of the system.

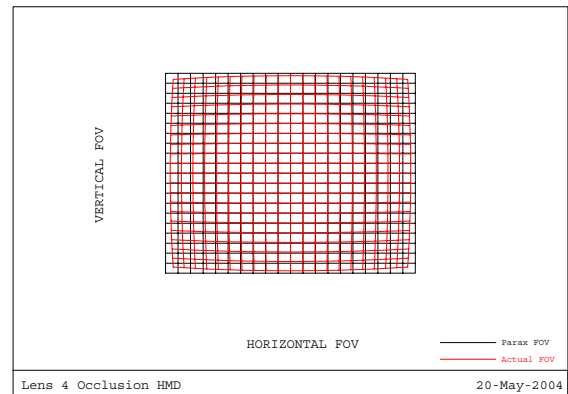


Figure 9. Distortion grid for the two element system

## 7. Experimental Implementation in the Laboratory

Before building custom optics, using the design presented in Section 4, we instrumented a prototype, shown in Fig. 10, with commercially available components to demonstrate feasibility of the optical



approach to occlusion. Listed in the order of light propagation path, starting at the left and propagating to the right, a white light source, a diffuser sheet, a transparency as the object, an achromatic lens (i.e. doublet), a Moxtek polarizing beam splitter, a liquid crystal optical shutter and an F-LCOS from CRL Opto connected to the driver electronics. Note that the F-LCOS was placed physically underneath the shutter. The user or a camera taking a picture of the modulated scene would look through the image formed by Lens 2. Lens 2 is acting as a weak magnifier.

The rationale for choosing the transparency instead of a view of the laboratory as seen through the optics is to increase the contrast ratio of the image formed by the objective. In a future prototype, we will design an enclosure for the system where the ambient light can be blocked. With an enclosure, the image formed by the objective will not compete with the ambient light within the laboratory and the laboratory scene can be used as the object instead of the transparency.

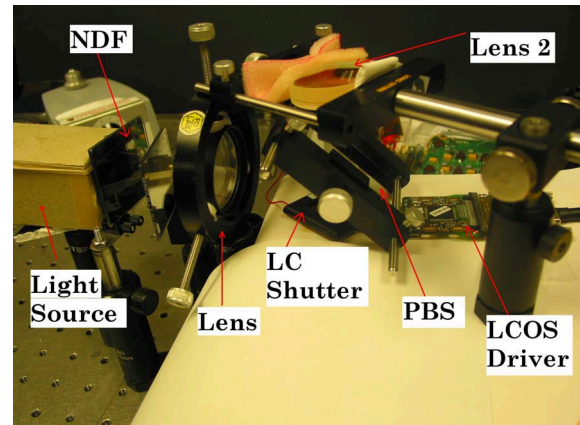
The achromat formed a fairly good image on the F-LCOS. It is clear however that a custom designed optical system as proposed in section 6 will provide an image quality to full specification across the full FOV. The quality of the image formed was acceptable to the human eye, while it was difficult to sharply focus on the image while taking a picture. The essential test is the light blocking and passing capability of the whole system.

Fig. 11 shows the original object, a Monet painting, that was used in the prototype setup. This image was copied from a book onto a transparency in color. We used the area around the child in the painting as the object. The transparency was fixed between a polarization preserving diffuser sheet from Microsharp and the light source. This diffusing sheet ensured that the object was illuminated uniformly, to the best of our ability within our preliminary setup. We used a white light source from American Optical. The object (i.e., Monet painting) was imaged through the achromatic lens of focal length of 50mm. The output of the lens went to a polarizing beam splitter, fixed at a 45 degree angle, then onto the liquid crystal shutter and finally to the F-LCOS.

We would like to highlight the nature of operation of our F-LCOS module since it has a direct impact on the results presented here. In addition, the nature of the F-LCOS operation will make clear the need for the liquid crystal shutter. The F-LCOS operates with a field sequential scheme[9].

For a typical application, this means cycling through the red, green and blue pixels. Each cycle consists of displaying one color (for example, red) and providing telecentric illumination that matches the

color displayed at that moment. Additionally, the F-LCOS uses a 50:50 drive scheme. This means that only half the time allocated for a pixel is displayed on the display/SLM and the other half is used by driving the display/SLM with the negative of the pixel. In our application we are interested in blocking/passing parts of the scene.



**Figure 10. Experimental setup on the optical bench.**

Therefore the system operates in a black and white mode. In a 50:50 drive scheme, the display will show a "white" pixel for half the pixel cycle and a "black" pixel for the rest of the cycle. A 50:50 drive scheme is required to DC balance the liquid crystal to avoid charge migration. The 50:50 drive scheme also implies that a shutter will be necessary to block the negative image. Due to the persistence of the human visual system, the positive and negative cycles would be integrated in the eye, leading to a very low contrast (1:2) image without the shutter.

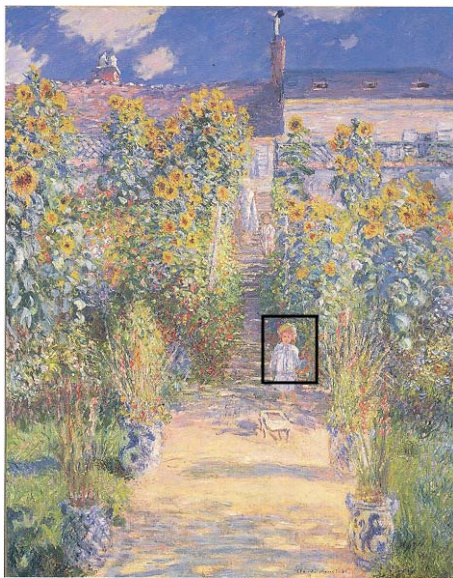
Depending on the orientation of the polarizers with respect to the F-LCOS module, one color will pass all (for example, white) the light and whereas the other color (for example, black) will block the light. The input light to the display must be linearly polarized. The simplest model for the F-LCOS is an electronically controllable quarter-wave plate laying on top of a perfectly flat mirror. If the pixel is turned off or "black" the orientation of the half-wave plate is aligned to the polarization, with no rotation of the polarization occurring. Thus the output light would be blocked by the analyzer. If the pixel is turned on or "white" the halfwave plate would be rotated at 45degrees (ideal) to the input polarization, and thus the resulting output polarization of the light would be at 90 degrees to the input. This light would then transmit through the analyzer. In practice, the display does not



reach the 45 degree level of performance, and thus the light is not fully rotated to 90 degree at the output, leading to a loss in throughput [12].

In order to reduce vignetting, we are using two Moxtek PBSs side by side. The image is modulated on the F-LCOS, changes polarization, is reflected back, and is transmitted through the beam splitter. This is our basic setup. However, in order to take a picture of this setup, we are using a second lens to image the output of the F-LCOS.

Fig. 12b is a photograph of the optical image as would be seen through the head-worn display, with no modulation (no occlusion) on the original scene. For comparison purposes, Fig. 12a is a photoshop scaled version of the region of interest shown in Fig. 11, therefore, it looks slightly pixelated. In the basic setup, we are imaging a relatively small field of view and also lens 2 is hardly magnifying the image. The significance of the result is that we can form an optical image of the scene on the F-LCOS and modulate it for occlusion.



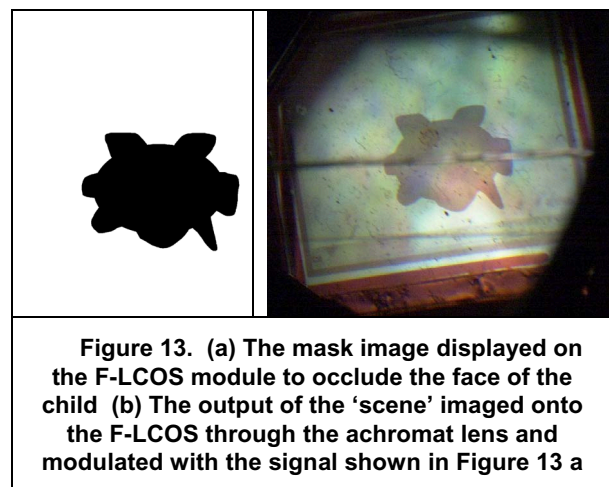
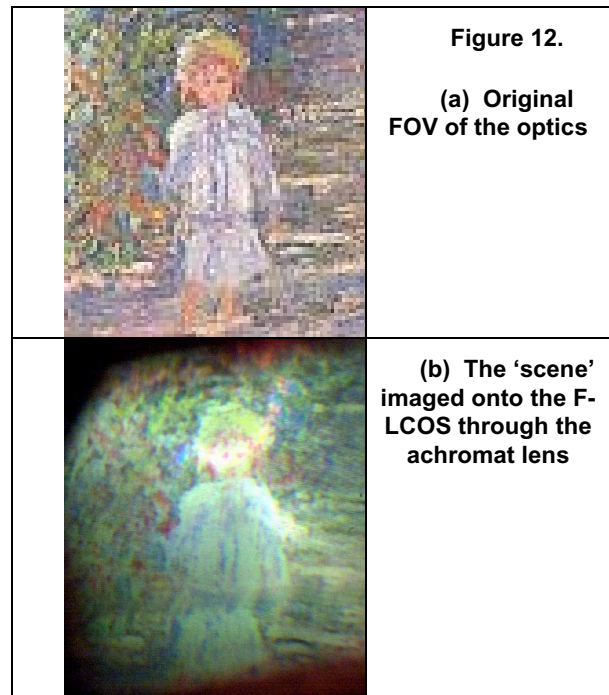
**Figure 11. The 'scene' used as an object while testing our head-worn display prototype (area around the child is used in the prototype, note the rectangle that indicates this area)**

Fig. 13a shows the mask signal that will modulate the scene. Fig. 13b shows the an image of the mask seen through the lens 2 on the F-LCOS superimposed on the image and in best focus we achieved (within the digital camera capability that we used to take the pictures). We can observe that head of the child is blocked according to the mask, which can have practically any shape and can be updated at video rates. This first result, which points to the promise of

this new technology, also points to the need to further work on the engineering aspects of the system to improve the contrast ratio of the mask which appears to be scene illumination dependent.

Furthermore, the line going across the middle of the image is the contact point of the polarizing beam splitters. The two PBSs were placed side by side to cope with the fact that we did not have a large enough beam splitter from the off the shelf components.

Also, importantly in assessing the results, the reason that the background appears blurry is because the mask perceptually currently appears in a different plane than the modulated image of the scene. This was confirmed using parallax.



Due to a couple of millimeters discrepancy between the image and the mask map centimeters in visual space after the eyepiece (i.e. lens2 in Fig. 10), the camera cannot at this time capture both in focus at the same time. So we focused the camera on the image first to get Fig. 12b, and on the mask next to get the occlusion image shown in Fig. 13b. The images shown in Fig. 12 and Fig 13 were taken with an  $f/22$  setting on the camera to try to increase depth of field to the maximum value available for the camera.

## 8. Conclusion

We presented a novel, compact, and optical approach to the occlusion problem. We provided a comprehensive description of the working of the system. We analyzed feasibility and early performance characteristics of the proposed system. Further, we presented preliminary results of both the lens design image quality and demonstrated the occlusion display based on a reflective modulator and polarization-based optics. In the current implementation, one limitation is the viewpoint offset which is non negligible in personal space ( $< 3\text{m}$ ) where the user is manipulating objects.

## 9. Future Work

Future work includes an advanced optimization of the lens design with a  $> 20\%$  MTF value at the maximum spatial frequency criteria. We will build custom optics based on an optimized version of the design presented in this paper. We will also investigate an optimized shutter mechanism to perform the occlusion in the same plane as the image, which we could not achieve using off-the-shelf optics. An immediate target for our system is to push the contrast ratio by using optimized custom optics that would be as close to  $f/2$  as possible. Also, we will take special care to illuminate the F-LCOS telecentrically. Optimized optics would have provisions to disallow ambient light. We also want to study the polarization issues in more depth. We believe, all these improvements would yield the advertised  $> 1:200$  contrast ratio of our SLM, thereby improving the current prototype presented in this paper. Finally, we plan to add depth sensing in order to have real objects occluding the virtual.

## 10. Acknowledgements

The authors would like to thank D. Vettese, P. Holmes, D. York and S. Clarke at CRL Opto for

excellent support. This work was supported by the Office of Naval Research grant N00014-03-10677. We would like to thank the anonymous reviewer who pointed out the viewpoint offset issue.

## 11. References

- [1] Rolland and Fuchs. Optical versus video see-through head mounted displays. In *Fundamentals of Wearable Computers and Augmented Reality*.
- [2] Kiyoshi Kiyokawa, Mark Billingham, Bruce Campbell, Eric Woods: An Occlusion-Capable Optical See-through Head Mount Display for Supporting Co-located Collaboration. *ISMAR 2003*: 133-141
- [3] US 6,037,914. Robinson. Method and apparatus for augmented reality using a see-through head-mounted display.
- [4] T. Uchida, K. Sato and S. Inokuchi. "An Optical See-through MR Display with Digital Micro-mirror Device", *Transactions of the Virtual Reality Society of Japan*, Vol. 7. No.2, 2002.
- [5] Kingslake. *History of the Photographic Lens*. Academic Press, 1989.
- [6] J. Melzer and K. Moffitt. *Head Mounted Displays: Designing for the User*. Mc-Graw Hill.
- [7] Ed. Benton. *Selected papers on three-dimensional displays*. SPIE Milestone Series, Volume MS 162.
- [8] Eds. Bass et.al. *Handbook of Optics*. Vol I and II.
- [9] CRL Opto SXGA Field Sequential Color Application Note. <http://www.crlopto.com>
- [10] CRL Opto. Application guide for ferroelectric liquid crystal on silicon displays. P. 7.
- [11] CRL Opto. User manual of the SXGA-R2 system.
- [12] Vettese. Personal communication. 15 March 2004.
- [13] Kocian, D.F., "Design considerations for virtual panoramic display (VPD) helmet systems," Armstrong Aerospace Medical Research Laboratory, Visual Display Systems Branch, Wright Patterson Airforce Base, Dayton, Ohio 45433-6573 (1988).
- [14] O. Bimber and B. Frochlich. Occlusion Shadows: Using Projected Light to Generate Realistic Occlusion Effects for View-Dependent Optical See-Through Displays. In *Proc. International Symposium on Mixed and Augmented Reality (ISMAR '02)*, pp.186-195, 2002.
- [15] Y. Ha, J. Rolland and O. Cakmakci. A compact Optical See-Through Head-Worn Display with Occlusion Support. US Patent filed, 2004.
- [16] E. Tatham. Getting the best of both real and virtual worlds. *Communications of the ACM*, 42: 9, p 96-98.