

A NOVEL SEE-THROUGH SCREEN BASED ON WEAVE FABRICS

Cha Zhang[†], Ruigang Yang[‡], Tim Large[†], Zhengyou Zhang[†]

[†]Microsoft Research, One Microsoft Way, Redmond, WA 98075
{chazhang,tlarge,zhang}@microsoft.com

[‡]University of Kentucky, 1 Quality Street, Lexington, KY 40507
ryang@cs.uky.edu

ABSTRACT

See-through screens (STS) have found important applications in remote collaboration systems to enhance non-verbal communication and gaze awareness. Existing STS designs often sacrifice the display quality significantly, rendering low-contrast images that discount the overall user experience. In this paper, we present a novel see-through screen solution based on weave fabrics. Such fabrics are known to be acoustically transparent and used to build professional projection screens for Hollywood studios. We place a camera immediately behind the screen and synchronize it with a 120Hz projector to perform time-multiplexing display and video capture. By focusing the camera at the user 4-5 feet away from the screen, the image of the weave fabric will be severely blurred. We present the imaging principle of the setup, and derive image processing techniques to enhance the quality of the captured video. The overall system is low cost, has much better display quality than existing systems, and can be used to build wall-size see-through screens for various applications.

Index Terms— See-through screen, gaze, remote collaboration

1. INTRODUCTION

As globalization continues to spread throughout the world economy, it is increasingly common to find projects where team members are widely distributed across continents. Videoconferencing has long been considered a critical technology to reduce high travel expenses for distributed workforces. Nevertheless, even with high end teleconferencing solutions such as HP's Halo system and Cisco's Telepresence system, face-to-face meeting is usually still a better experience than remote meetings.

One of the factors that are known to be essential for face-to-face communication is eye contact. As Simmel remarked [1], eye contact "represents the most perfect reciprocity in the entire field of human relationship". It instills trust and

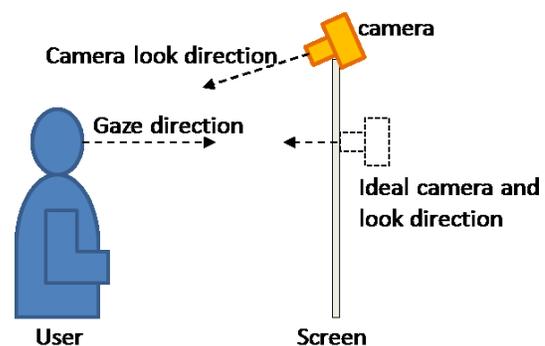


Fig. 1. The eye contact issue in typical videoconferencing systems. The ideal camera shall point to the user from behind the screen.

fosters an environment of collaboration and partnership. Lack of eye contact, on the other hand, may generate feelings of distrust and discomfort. Unfortunately, eye contact is usually not preserved in typical videoconferencing systems. As shown in Fig. 1, the user often looks at the remote party displayed on the screen, while the local camera often captures the user from above the screen, creating gaze disparity. To preserve eye contact, the ideal camera shall be placed behind the screen.

Creating a see-through screen has attracted research studies for decades. Early approaches such as Teleprompter [2] and Gazecam [3] used half silvered mirrors. Since half silvered mirrors need to be angled by around 45 degrees from the display, the footprint of such systems are usually inconveniently large. In addition, stray reflections caused by the mirror may also be distracting. The Clearboard system [4] designed a clever "Drafter-Mirror" architecture which improved upon simple half silvered mirror implementations. They placed the display surface at a 45 degree angle with respect to the ground, and further adopted polarization films on the screen and the camera to avoid the displayed image being captured by the camera. However, polarization films cut down lights significantly, leading to poor display quality.

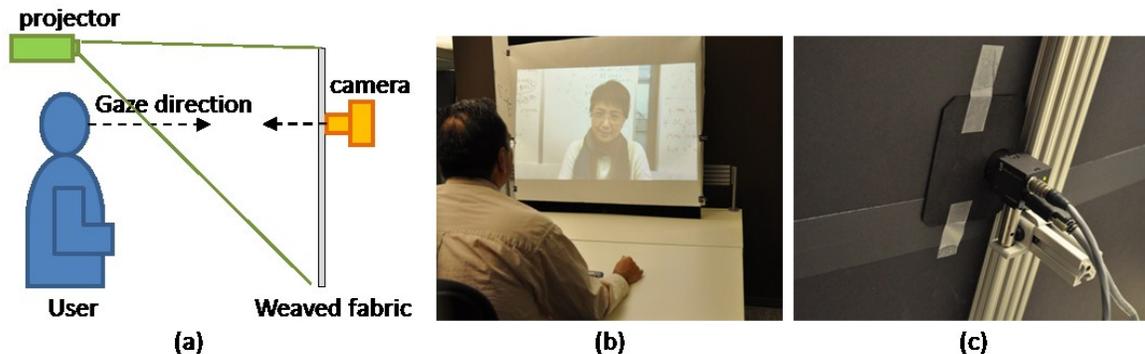


Fig. 2. The overall system setup. (a) System configuration. (b) Front side of the system with an ongoing teleconferencing. (c) The camera hidden behind the screen.

Another popular scheme to solve the eye contact issue is through switchable liquid crystal diffusers (SLCD), such as the system demonstrated by Shiwa and Ishibashi at NTT [5] and blue-c [6]. SLCD can switch between transparent and diffusing states based on the voltage applied on the film. When the diffuser, the camera and the projector are all synchronized, during the transparent state, the camera behind the diffuser can take images; during the diffusing state, the projector can render images of the remote user. A known limitation of SLCD is the speed of state switching that limits the update rate of the rendering, causing flickering images. In addition, the rendered images of SLCD tend to have poor contrast under ambient lighting conditions, making it unsuitable for many applications.

Recently holographic projection screens (DNP HoloScreen) have received a lot of attention [7][8][9]. These screens diffuse only light incident from pre-specified angles, and allow light to pass through otherwise. They do not require special synchronized cameras and projectors, thus offering greater freedom for designers. However, HoloScreen usually has severe backscatters, which shall be handled carefully. The TouchLight system [7] used infrared cameras to avoid difficulties, which is not suitable for teleconferencing. HoloPort [8] proposed to synchronize the projector and the camera to counter the backscatter problem. In the recent ConnectBoard system [9], Tan *et al.* proposed to use wavelength multiplexing to remove the backscatter. Additional processing is needed to pre-distort the projected image and color-correct the captured image in order to capture satisfactory images. Besides backscatters, Holo-Screen only diffuses part of the incident light from the pre-specified angle, and a crisp image can also be seen on the ceiling or the floor of the room. This means there are many angles for which the user is looking right into the beam, which can be quite painful. The display image quality of HoloScreen is similar to SLCD, which leaves room for improvement.

Another work that is closely related to ours is the MAJIC system by Okada *et al.* [10]. MAJIC adopted a screen that is a thin transparent film with a large number of small hexagons printed on both sides. The front side is printed with

white hexagons, thus it can be used as the projection screen. The back side has black hexagons, and one can see the other side through the screen. The MAJIC prototype used a 40% transmissibility screen, thus the displayed images had very poor quality, both in brightness and resolution.

In this paper, we present a novel see-through screen design by using *weave fabrics*. Such fabrics are known to be an excellent projection surface, and are widely used in professional movie studios for high display quality and acoustical transparency. We place a video camera right behind the screen, and capture videos by seeing through the small holes in the weave fabrics. We then design image processing algorithms to enhance the quality of the captured video. This results in a low-cost, high display quality see-through screen that can be used in many applications, such as teleconferencing, virtual reality, public advertisement, etc.

The rest of the paper is organized as follows. The hardware setup is explained in Section 2. The imaging model of the camera behind the screen is presented in Section 3. Video processing based on the imaging model and some experimental results are detailed in Section 4. Conclusions and future work are given in Section 5.

2. HARDWARE

The configuration of the proposed see-through screen is shown in Fig. 2 (a). We adopt a frontal projection scheme. The user is in the projection space, making it more space efficient for large systems. The camera is placed behind the screen (Fig. 2 (c)). As shown in Section 3 and 4, it is advantageous to place the camera as close to the screen as possible. The projector and the camera are time-synchronized using the V-sync signal of the projector input. For every other frame the projector will output a black image, allowing the camera to open its shutter to capture an image. To avoid flickering, we use a 120 Hz projector (DepthQ HD 3D projector), effectively refreshing the displayed image at 60 Hz.

The camera is a Flea3 FL3-FW-03S3C made by Point-Grey Research Inc. It is capable of capturing 640×480 pixel

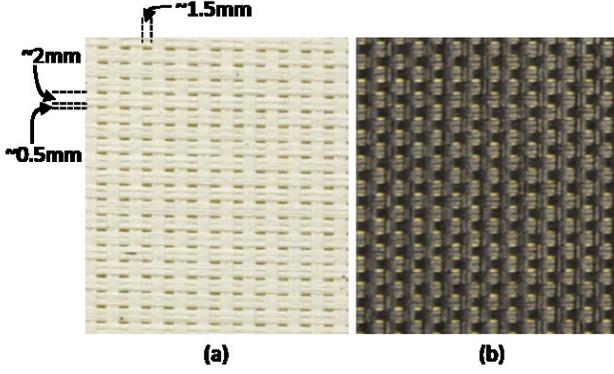


Fig. 3. The weave fabric screen. (a) Texture of the Phifer SheerWeave solar shade screen with 10% openness. (b) The back side of the screen is painted black to prevent backscatters.

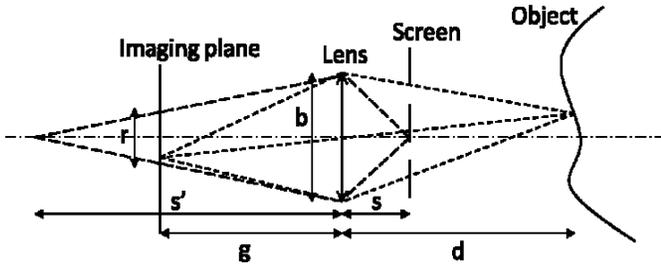


Fig. 4. Use thin lens model to analyze the blurring radius of the weave fabric screen.

images at 76 frames per second. Since the holes on the weave fabric screen are very small, we tested two lenses with large apertures, the Fujinon DF6HA-1B 6mm $f/1.2$ lens and the Pentax 12mm $f/1.2$ lens.

The weave fabric screen is obviously the most important component of our system. We started with samples of the CineWeave HD screen from SMX Cinema Solutions [11], which has about 5% openness. Later we found low cost alternatives by using the SheerWeave solar shade made by Phifer. We chose the 2100 P02 white shade with 10% openness as our screen for a good compromise between display quality and see-through video quality. The texture of the screen is shown in Fig. 3. We painted the back side of the screen black to prevent backscatters.

3. IMAGING MODEL

We first apply the thin lens model to analyze the blurring diameter of the defocused weave fabric, as shown in Fig. 4. Let the focal length of the lens be f . Assume the camera focuses on the object at distance d with the imaging plane at g . According to the thin lens equation, we have:

$$\frac{1}{d} + \frac{1}{g} = \frac{1}{f}. \quad (1)$$

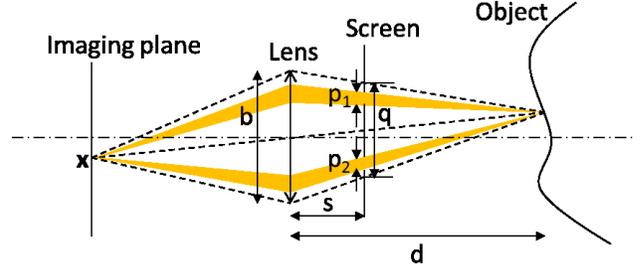


Fig. 5. Illustration of the see-through ratio $\alpha(\mathbf{x})$ in our system.

If the screen is at distance s , it will be focused at s' , where:

$$\frac{1}{s} + \frac{1}{s'} = \frac{1}{f}. \quad (2)$$

The blurring diameter r satisfies:

$$\frac{s' - g}{s'} = \frac{r}{b}, \quad (3)$$

where b is the aperture of the lens. Consequently, we obtain:

$$r = \frac{f^2(d-s)}{Ns(d-f)}, \quad (4)$$

where $N = f/b$ is the f-number of the lens. It can be seen from Eq. (4) that increasing the focal length f , reducing the f-number N and decreasing the screen distance s will all enlarge the blurring diameter r .

We next present an idealized model for the image formation process. Let the captured image be denoted as $I(\mathbf{x})$, where \mathbf{x} is the pixel index. Inspired by Aydin and Akgul [12], we represent:

$$I(\mathbf{x}) = [\alpha(\mathbf{x})L(\mathbf{x}) + (L_s * H_r)(\mathbf{x})]v(\mathbf{x}) + \epsilon(\mathbf{x}), \quad (5)$$

where $L(\mathbf{x})$ is the object radiance observed by pixel \mathbf{x} when there is no screen occlusion. $\alpha(\mathbf{x})$ is the see-through ratio, which will be detailed later. L_s is the radiance of the back side of the screen, and H_r is the blurring filter with diameter r , which is given by Eq. (4). Since we have painted the back side of the screen black, $L_s \approx 0$. Therefore, the term $(L_s * H_r)(\mathbf{x})$ can be safely ignored in our application. $v(\mathbf{x})$ is used to model scene independent effects such as vignetting, which happens often in low-cost lenses. $\epsilon(\mathbf{x})$ is the sensor noise.

Fig. 5 illustrates the computation of the see-through ratio $\alpha(\mathbf{x})$. Since the object surface can only be seen through the openings $p_i, i = 1, 2, \dots$, we have:

$$\alpha(\mathbf{x}) = \frac{\sum_i A_{p_i}}{A_q}, \quad (6)$$

where A_{p_i} is the opening area of hole i , and A_q is the overall area of lights at the screen distance that enter the lens and reach \mathbf{x} when there is no occlusion.



Fig. 6. Images of a pure white paper in front of the lens. (a) 6mm lens. (b) 12mm lens.

Let A_b denote the area of the aperture, since

$$\frac{A_q}{A_b} = \frac{(d-s)^2}{d^2}, \quad (7)$$

if the object is at a constant depth, A_q will be a constant proportional to the aperture area A_b . However, since the holes of the weave fabric have finite sizes, different openings will be seen when varying \mathbf{x} . Hence $\alpha(\mathbf{x})$ varies depending on \mathbf{x} . For a piece of weave fabric with average openness α_0 ,

$$\alpha(\mathbf{x}) \rightarrow \alpha_0, \text{ if } \max_i \frac{A_{p_i}}{A_q} \rightarrow 0. \quad (8)$$

In other words, to reduce the fluctuation of the see-through ratio, we either increase the lens aperture, or reduce the size of the holes in the fabric. In Section 4.1, we will further discuss how to correct the unevenness of the see-through ratio digitally.

For a real-world lens, the see-through ratio may be affected by other factors, such as lens distortions, edge effects, etc. The weave fabrics also have unevenness in the size and distribution of holes. Fig. 6 (a) and (b) shows two images of a pure white paper ($L(\mathbf{x}) = L_0$ is a constant) captured by 6mm and 12mm lenses, respectively. Note the see-through ratio of the 12mm lens fluctuates much less than that of the 6mm lens. This is consistent with Eq. (8), since the 12mm lens has a larger aperture.

4. VIDEO PROCESSING

Due to the small openness of the weave fabric used in our system and the short camera exposure time (7ms per frame), the captured images are often dark and noisy (see Fig. 7 (a) and Fig. 8 (a) for some examples). In this section, we present a two-step video enhancement process to improve the video quality: recovering the object radiance and video denoising.

4.1. Recovering the Object Radiance

In Section 3, we presented an imaging model for the camera seeing through the weave fabric. Although the see-through



Fig. 7. Recover the object radiance for video enhancement. (a) Raw image captured by the camera. (b) After radiance recovery using Eq. (12). Note no further image enhancement schemes such as histogram equalization were used to produce these images.

ratio $\alpha(\mathbf{x})$ is hard to predict, we can still recover the object radiance as follows.

We first capture a video of a static, pure white object (e.g., a white paper), as was shown in Fig. 6. According to Eq. (5), the received images can be represented as:

$$I(\mathbf{x}) = \alpha(\mathbf{x})v(\mathbf{x})L_0 + \epsilon(\mathbf{x}), \quad (9)$$

where L_0 is the constant radiance of the white object. Note the term $(L_s * H_r)(\mathbf{x})$ in Eq. (5) has been ignored since $L_s \approx 0$.

By averaging these video frames to obtain a mean image $\bar{I}_0(\mathbf{x})$, we effectively remove the sensor noises, obtaining:

$$\bar{I}_0(\mathbf{x}) = \alpha(\mathbf{x})v(\mathbf{x})L_0. \quad (10)$$

That is, the mean image $\bar{I}_0(\mathbf{x})$ captures the fluctuation of $\alpha(\mathbf{x})$ and scene independent effects $v(\mathbf{x})$. For an arbitrary scene captured by the same camera, we have:

$$I(\mathbf{x}) = \alpha(\mathbf{x})v(\mathbf{x})L(\mathbf{x}) + \epsilon(\mathbf{x}). \quad (11)$$

Combining Eq. (10) and (11), we have:

$$L(\mathbf{x}) \approx \frac{I(\mathbf{x})L_0}{\bar{I}_0(\mathbf{x})}, \text{ when } \bar{I}_0(\mathbf{x}) \gg \epsilon(\mathbf{x}), \quad (12)$$

which can be computed very efficiently for each frame. The condition $\bar{I}_0(\mathbf{x}) \gg \epsilon(\mathbf{x})$ is usually satisfied if a large aperture lens is placed right behind the screen and focused on objects a few feet away from the screen.

Fig. 7 shows some results for object radiance recovery using Eq. (12). It can be seen that the images are improved significantly compared with the original images captured by the camera.

It is worth mentioning that if the aperture of the lens is very large or the weave fabric holes are very small (assume the same openness), $\alpha(\mathbf{x})$ will be near constant, and the raw image captured by the camera will be a very good approximation of the radiance after typical scene independent processes such as vignetting removal.

4.2. Video Denoising

One may notice from Fig. 7 that the images captured by the camera are very noisy. These noises remain in the processed images after radiance recovery. It is necessary to perform video denoising to further improve the video quality.

Video denoising has been an active research topic for many decades. A few novel and effective video denoising approaches have been proposed recently, such as non-local means [13], wavelet domain filters [14], bilateral filters [15], etc. Unfortunately, most of these algorithms are very computationally expensive. In our system, the videos are captured at 60 frames per second. In order to perform real-time video denoising, we adopted the simple scheme of temporal denoising. The optical flow of neighboring frames is first estimated using the well-known Lucas-Kanade method [16]. Corresponding pixels based on the optical flow are then averaged to obtain the denoised image. The algorithm is implemented with multiple threads, each handling a sub-region of the input video frame. In the future we plan to look into GPU-based implementations to further speed up the denoising process.

Fig. 8 shows two examples of video denoising results. The images are cropped from the original 640×480 images to demonstrate the denoising effect. It can be seen clearly that the algorithm is very effective in removing sensor noises.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we presented a novel see-through screen based on weave fabrics. It has better display quality than most of the existing see-through screens. The see-through video quality is also very good, thanks to our radiance recovery algorithm based on the camera’s imaging model and real-time video denoising. The screen can be used for teleconferencing systems that maintain ideal eye contact between attendees. Since weave fabrics are manufactured with mature technology, they can be used to build low-cost, wall-size screens with multiple see-through cameras.

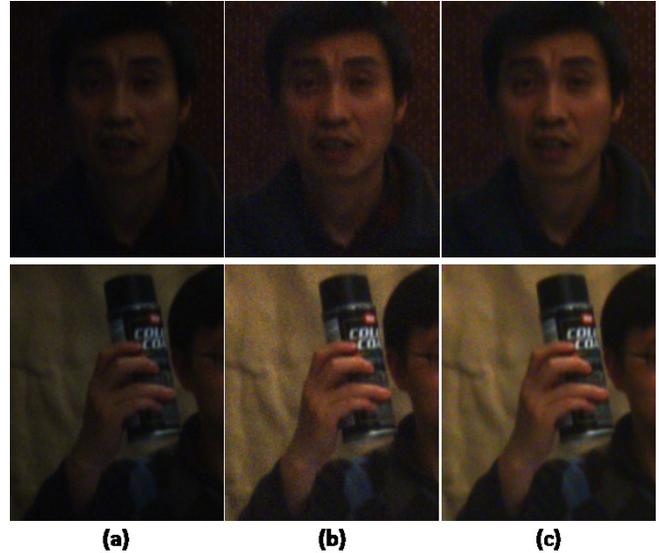


Fig. 8. Video denoising results. (a) Raw image captured by the camera (cropped). (b) After radiance recovery using Eq. (12). (c) After noise removal.

When designing see-through screens, there has always been a tradeoff between the display quality and the see-through video quality. For weave fabrics, the openness of the screen is a key factor determining the tradeoff. Fortunately, it is very easy to change the openness of weave fabrics during manufacture. We have tested with fabrics with 10% openness, though the optimal openness is still to be explored.

In addition, due to the small openness of weave fabrics, the aperture of the camera behind the screen is usually large in order to receive sufficient lights and reduce the adverse impact of screen occlusion. This can reduce the depth of field of the camera. An interesting future direction is to enable adaptive focusing, such that the camera can always focus on the objects in front of the screen.

REFERENCES

- [1] G. Simmel, “Sociology of the Senses: Visual Interaction,” in R. Park and E. Burgess, editors, *Introduction to the Science of Sociology*, University of Chicago Press, 1921.
- [2] J. Oppenheimer, “Prompting apparatus,” US Patent 2883902, 1959.
- [3] S. Acker and S. Levitt, “Designing videoconference facilities for improved eye contact,” *Journal of Broadcasting and Electronic Media*, Vol. 31, No. 2, pp. 181–191, 1987.
- [4] H. Ishii and M. Kobayashi, “Clearboard: a seamless medium for shared drawing and conversation with eye contact,” *ACM SIGCHI* 1992.

- [5] S. Shiwa and M. Ishibashi, "A large-screen visual telecommunication device enabling eye contact," *SID Digest*, Vol. 22, pp. 327–328, 1991.
- [6] M. Gross, S. Wurmlin, M. Naef, E. Lamboray, C. Spagno, A. Kunz, E. Koller-Meier, T. Svoboda, L. Van Gool, S. Lang, K. Strehlke, A. V. Moere, and O. Staadt, "Blue-c: a spatially immersive display and 3d video portal for telepresence," *ACM SIGGRAPH* 2003.
- [7] A. Wilson, "Touchlight: An imaging touch screen and display for gesture-based interaction," *ICMI* 2004.
- [8] M. Kuechler and A. Kunz, "Holoport - a device for simultaneous video and data conferencing featuring gaze awareness," *ICVR* 2006.
- [9] K. Tan, I. Robinson, R. Samadani, B. Lee, D. Gelb, A. Vorbau, B. Culbertson and J. Apostolopoulos, "ConnectBoard: a remote collaboration system that supports gaze-aware interaction and sharing," *MMSP* 2009.
- [10] K. Okada, F. Maeda, Y. Ichikawaa and Y. Matsushita, "Multiparty videoconferencing at virtual social distance: MAJIC design," *CSCW* 1994.
- [11] SMX Cinema Solutions, <http://www.smxscreen.com/>.
- [12] T. Aydin and Y. Akgul, "An occlusion insensitive adaptive focus measurement method," *Optics Express*, Vol. 18, No. 13, June 2010.
- [13] A. Buades, B. Coll and J.-M. Morel, "Nonlocal image and movie denoising," *International Journal of Computer Vision*, Vol. 76, No. 2, pp. 123-139, 2008.
- [14] I. Selesnick and K. Li, "Video denoising using 2D and 3D dual-tree complex wavelet transforms," *Wavelets: Applications in Signal and Image Processing X*, SPIE, Vol. 5207, pp. 607-618, 2003.
- [15] M. Zhang and B. Gunturk, "Multiresolution bilateral filtering for image denoising," *IEEE Trans. on Image Processing*, Vol. 17, No. 12, pp. 2324-2333, 2008.
- [16] B. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," *Proc. of DARPA Image Understanding Workshop*, 1981.