

Transparent Stereoscopic Display and Application

Nicola Ranieri, Hagen Seifert, Markus Gross

Computer Graphics Laboratory, ETH Zürich, www.graphics.ethz.ch

ABSTRACT

Augmented reality has become important to our society as it can enrich the actual world with virtual information. Transparent screens offer one possibility to overlay rendered scenes with the environment, acting both as display and window. In this work, we review existing transparent back-projection screens for the use with active and passive stereo. Advantages and limitations are described and, based on these insights, a passive stereoscopic system using an anisotropic back-projection foil is proposed. To increase realism, we adapt rendered content to the viewer's position using a Kinect tracking system, which adds motion parallax to the binocular cues. A technique well known in control engineering is used to decrease latency and increase frequency of the tracker. Our transparent stereoscopic display prototype provides immersive viewing experience and is suitable for many augmented reality applications.

Keywords: 3-D display, transparent, back-projection, stereoscopic display, eye tracking, motion parallax

1. INTRODUCTION

In our society of the digital age, information has become one of the most important resources. Data is acquired anytime and everywhere, processed, stored or played back to users, who often are over-strained by the mass of information. With augmented reality (AR), efforts have been made to embed such information into the real world to make them easier accessible and readable.

Example given, head mounted displays (HMD) can overlay an image with a user's field of view, providing additional information to what he is looking at. Prominent example is the Google Glass [1], but many others exist. Although being compactly build, the additional hardware required on the glasses often makes them cumbersome to wear.

For a different kind of applications, Samsung uses a light-box with a transparent liquid crystal display (LCD) in front of it to show additional information for a product exposed in the box [2]. Similarly, LG uses a transparent LCD as door of a fridge to inform a user about content while giving clear sight of the food inside [3]. However, transparent LCDs are not self-emissive, transmit only 33% of the light through the color filters and absorb another 50% in the polarizer. Hence, extremely bright background is needed to account for the resulting transparency of usually less than 17%, which narrows the area of application. Transparent OLED, a promising alternative, are self-emissive but also could not achieve sufficient transparency so far.

Sun Innovations propose a self-emissive solution with high transparency: A fluorescent film displays content when activated by a galvanomic blue-ray laser scanner [4]. The screen can be used in environments with any brightness, but so far it has been used to show 2D content only.

Few attempts have been made to use transparent display technology for visualizing 3D content. Lee and Hong use a semi-transparent concave mirror array to create transparent 3D [5]. Mock et al. use a back-projection foil with active shutter glasses in a tele-collaborative environment, providing high transparency [6]. In a similar system presented by Kuster and colleagues, a remotely located person is rendered in 3D on a transparent screen for immersive telecommunication [7].

One problem of such transparent 3D systems is how to align virtual content with the real world. This can be achieved by tracking a viewer's eye position relative to the screen and adapting the rendered content accordingly. Though, tracking devices usually suffer from low refresh rates and high latency.

These issues can be overcome using a Kalman filter and prediction, which has been successfully done for human motion and opaque 3D screens in previous work [8]. But transparent 3D displays demand a much higher accuracy and lower latency, as the viewer can see and compare real objects and virtual content at the same time with respect to their movement. Thus, these existing approaches have to be re-evaluated to see whether the quality still can achieve an immersive experience.

A comprehensive review of recent development in rendering, tracking and display technology for augmented reality is given by Zhou et al. [9].

In this work, we assess both isotropic and anisotropic transparent back-projection technologies for the use in combination with passive and active shutter glasses. Based on the derived insights, we propose and build a transparent stereoscopic 3D system for augmented reality. Next, we apply Kalman filters and prediction to a Kinect tracker to verify that it can be used to provide proper perspective cues and motion parallax on a transparent stereoscopic screen. Results and a discussion conclude our work.

2. BACK-PROJECTION SCREENS

In applications with a static screen and sufficient space for a projector setup, transparent back-projection screens offer a simple and convenient way to overlay virtual 3D content with the real environment. Generally speaking, there are two kinds of back-projection technologies: anisotropic and isotropic ones. Anisotropic back-projection screens redirect or diffuse light only if coming from a specific direction, where isotropic back-projection screens diffuse all incoming light from any direction. Both can be used either with active or passive 3D glasses. For active shutter glasses, consecutive frames show views for the left and right eyes, which are then blocked by the left or right shutter glass accordingly. For passive glasses, different polarizations are used for the left and right view, which are filtered in the corresponding polarizer of the glasses.

Anisotropic Back-Projection Screens selectively redirect light coming from a specific direction [10][11]. Usually they are optimized such that for each point on the screen only the light rays from a specific center of projection are first redirected to coincide with the screen surface normal and then spread to form a certain field of view. Hence, this approach is very robust against environment light, as incident light from other directions than the optimized center of projection can pass unhampered. As drawback, the setup is fixed and must maintain distance as well as direction between screen and projector. Furthermore, much light passes these screens without being redirected or diffused and hence the approach is not very light efficient.

Anisotropic back-projection screens can be realized with diffractive optical elements (DOE): A holographic grating is used to redirect and shape the incident beams. Therefore such screens also suffer from the rainbow effect, as different wavelengths are redirected differently.

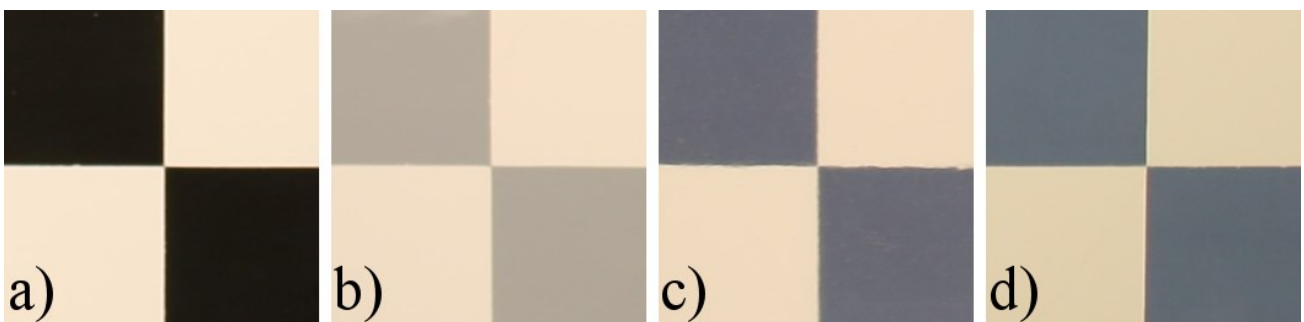


Figure 1: Comparison of different back-projection screens. a) Shows a test pattern with no screen in front used as ground truth. b) Illustrates the test pattern with an isotropic back-projection screen [12] in front. c) Shows the effect of an anisotropic foil [10] and d) the effect of an anisotropic glass [10]. The low contrast common to isotropic screens is clearly visible when comparing a) and b). Anisotropic glass and foil perform equally well with the foil being a little brighter than the glass.

The holographic film itself is usually polarization preserving and hence, both passive and active stereo can be used with such screens. However, the coating or the material in which the film is embedded might be depolarizing and thus this property has to be verified by the manufacturer. Due to the structure in the DOE, the polarizers for the left and right view do not only have to be aligned with the filters in the glasses but also towards the screen. In addition, crosstalk is higher at flat viewing angles, narrowing the field of view. The active approach mostly suffers from the known drawbacks as bulky glasses and required line of sight to the synchronization beacon.

Isotropic Back-Projection Screens diffuse incident light from all angles into all directions [12]. Hence the projectors can be positioned anywhere towards the screen, giving the setup much flexibility. On the other hand it is weak against environment light as also light e.g. from a lamp will be diffused, lowering the quality of the displayed image.

Isotropic back-projection screens can be realized by particles or droplets embedded in a clear material. Light intersecting with these droplets are diffused into all directions. As light is diffused even within the screen and also is totally internally reflected, subsurface scattering occurs. This creates a blurred ghost around the displayed image. Also, the ratio between area covered by droplets and clear area defines the transparency as well as the brightness of displayed images. To be able to compete with the brightness of other technologies, this trade-off usually is chosen in a way which is resulting in poor contrast. A comparison between isotropic and anisotropic screens is illustrated in Figure 1.

Isotropic back-projection foils are sold as polarization preserving which is true for the directly diffused light. However, the light which is scattered screen internally loses its polarization and is hence perceived as crosstalk in a passive stereo approach, where it is perceived as low frequency/blurred halo with active shutter glasses.

Table 1 gives a summarized overview and lists the advantages and drawbacks of each approach in use with active or passive stereo.

	anisotropic	isotropic
passive	+ cheap lightweight glasses	+ cheap lightweight glasses
	- light inefficient	- poor contrast
	+ robust against environment light	- weak against environment light
	- fix projector position	+ arbitrary projector position
	- crosstalk at flat viewing angles	- crosstalk at flat viewing angles
	- rainbow effect	- subsurface scattering perceived as crosstalk
active	- heavy glasses and IR beacon	- heavy glasses and IR beacon
	- light inefficient	- poor contrast
	+ robust against environment light	- weak against environment light
	- fix projector position	+ arbitrary projector position
	- rainbow effect	- subsurface scattering perceived as halo

Table 1: This table shows a comparison of anisotropic and isotropic back-projection screens in use with active or passive glasses.

3. HEAD TRACKING

To adapt the rendering perspective and to provide motion parallax to the user, head tracking is performed using the Microsoft Kinect. The Kinect was released by Microsoft as a motion sensing input device for the Xbox 360 in 2010. The subsequent release of Windows drivers and SDK, low cost and high availability make it interesting for research purposes as well. To overcome the Kinect's inherent latency and low frame rate, we apply motion prediction using a Kalman filter.

The sensor sends a color and a depth stream from its cameras to the host computer, which is used for tracking. The Kinect SDK and Developer Toolkit provide a set of useful libraries to process these data.

Skeleton Tracking provides the 3D positions of 21 joints of the tracked body. It is robust against arbitrary body orientation and even against partial occlusion. However, the head is represented by only a single joint and its accuracy lies only in the centimeter range.

Face Tracking detects faces with high accuracy and returns 121 3D vertices for the face incorporating the user specific face shape and current mimics. Unfortunately the high accuracy comes at the cost of robustness. The face is often lost when it is tilted or rotated away from the Kinect too much, or if the distance between face and sensor exceeds $1.5m$.

Sudden losses of tracked faces greatly disrupt the immersion and thus, continuity is of high importance. Hence, *Skeleton Tracking* is the preferred tracking method, despite its shortcomings in accuracy. Alternatively, both approaches can be combined, such that face tracking is used when available and skeleton tracking used when the face cannot be detected.

A common problem in either tracking method is the delay imposed by the Kinect sensor which is reported to be about $125ms$ [13]. If a user located $1m$ in front of the screen moves sideways at $1m/s$ and the virtual 3D content is $1m$ behind the screen, this content appears to be offset by $12.5cm$ in the opposite moving direction of the user. When the user stops his motion, content aligns in the following $125ms$, inhibiting the immersive experience. Furthermore, the sensor's frame rate is only 30 frames per second. This adds a varying delay between 0 and $33ms$ which leads to a perceived jitter of the content of up to $3cm$ using the same settings as above. Both the jitter and the delay can be counteracted by motion prediction using a Kalman filter.

4. MOTION PREDICTION

The Kalman filter is an algorithm used for stochastic state estimation from noisy sensor measurements and can also be used for motion prediction. It is used predominantly in military and civilian navigation systems such as GPS. More significantly for this work, it is also used extensively for tracking in interactive computer graphics [14].

The state of the tracked head is described by the following discrete time model:

$$x_k = Ax_{k-1} + w_{k-1} \quad (1)$$

x represents the state vector, A the state transition matrix and w the process noise which is assumed to be Gaussian white noise. The state vector contains the position, velocity and, optionally, the acceleration of the tracked feature. For position and velocity, the model can be found from Newton's laws of motion:

$$x = \begin{bmatrix} s \\ \dot{s} \end{bmatrix} \quad A = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix} \quad (2)$$

Measured values z relate to the estimated state by the observation model H and the normally distributed measurement noise v :

$$z_k = Hx_k + v_k \quad H = [1 \quad 0] \quad (3)$$

For the Kalman filter, the process and measurement noise covariance Q and R are required additionally, which are design parameters of the filter.

$$Q = E[w_k w_k^T] = \begin{bmatrix} \frac{1}{4} \Delta t^4 & \frac{1}{2} \Delta t^3 \\ \frac{1}{2} \Delta t^3 & \Delta t^2 \end{bmatrix} \sigma_a^2 \quad R = E[v_k v_k^T] = \sigma_m^2 \quad (4)$$

The execution can be divided into two steps. In the *prediction* step, the current state is estimated by propagating the previous one.

$$\hat{x}_k^- = A\hat{x}_{k-1} \quad (5)$$

$$P_k^- = AP_{k-1}A^T + Q \quad (6)$$

In the *correction* step, the resulting *a priori* estimate \hat{x}_k^- is updated with the measurement results to receive the improved *a posteriori* estimate \hat{x}_k .

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (7)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H \hat{x}_k^-) \quad (8)$$

$$P_k = (I - K_k H) P_k^- \quad (9)$$

A more detailed explanation of the Kalman algorithm and notation can be found in [14].

Multiple prediction steps are performed for each received Kinect frame, increasing the refresh rate and thus lowering the jitter caused by the low Kinect frame rate. To improve the overall latency of $125ms$, a further prediction by a constant amount of time can be applied.

The prediction results from the Kalman filter largely depend on the choice of Q and R . Furthermore, larger prediction times lead to noisy results and overshoot. In our experiments, prediction times of $67ms$ led to the best compromise in latency reduction and tracking quality, halving the Kinect's latency. The lowered latency and jiggling demonstrate the potential of motion prediction for the Kinect and for head tracking.

5. SYSTEM DESIGN

Design Choices: In this work we focus on an immersive screen adapting its content to the viewer position. Hence, it is only for a single user, who is looking at the screen mostly from a steep viewing angle. Thus we decided to use an anisotropic back-projection screen with passive glasses, as it provides most benefits and acceptable drawbacks with respect to our needs. The linear polarization glasses are lightweight and provide unencumbered viewing experience. The screen is robust against environment light and achieves high quality imaging, allowing immersion into the virtual reality. The slight crosstalk and rainbow effect are acceptable for our use cases.

Hardware: Our screen is a HOPS® Glass from Visionoptics [10]. We use two Acer H5360 projectors with linear polarizer to overlay two FullHD images on the projection screen. Distance between center of projection and center of the screen is $1.5m$ and the angle between screen normal and direction of the projection is 38° , as specified by the manufacturer. The DualHead2Go multi-display adapter from Matrox is used to provide the projectors with perfectly synchronized images. A Kinect is attached to the screen to track a viewer within the field of view of the display. The whole setup is shown in Figure 2.



Figure 2: Results captured on our prototype. Three different viewing positions are shown in the three images. Left and right images are filtered with the corresponding linear polarizer to simulate the effect of polarized glasses in addition to the change of perspective. The center image was captured without any filter and thus both views are overlaid. The three images together illustrate image separation and crosstalk as well as motion parallax.

Calibration: Images of a stereoscopic display must perfectly coincide to avoid eye strain and fatigue of the viewer. We correct the keystone and achieve rectification of the two images by calibrating the two projector coordinate spaces using homographies. A checkerboard pattern is displayed on each projector and captured from an external camera without changing camera position. Using four point correspondences between projector and camera coordinates, homographies can be computed which are then used to relate one projector image to the other.

6. RESULTS AND LIMITATIONS

Figure 2 shows three perspectives of a synthetic scene rendered on our physical prototype. The left image shows the view of the left eye from the left side. Similarly, the right image shows the view of the right eye from the other side. The images were captured on a Canon EOS-1D Mark III with a linear polarizer in front to simulate the effect of the polarized glasses. Both images show as well the image separation and crosstalk as the motion parallax when moving to different positions. The center image shows a center position with no filter in front of the camera. Hence, both views are overlaid and the captured image is brighter. The rainbow effect, common to anisotropic back-projection screens, is not visible in the photos and only barely visible by eye.

Quality of the improved tracking results are shown in Figure 3. The plots compare predicted, up-sampled and filtered signals with different prediction time (dashed lines) with the measured signal (dotted line). The predicted signal decreases in quality with increasing prediction time. Our system is able to provide motion parallax for slow viewer motion. If the viewer moves fast, further improvements or a better tracker would be required, as the prediction time of $67ms$ becomes too large.

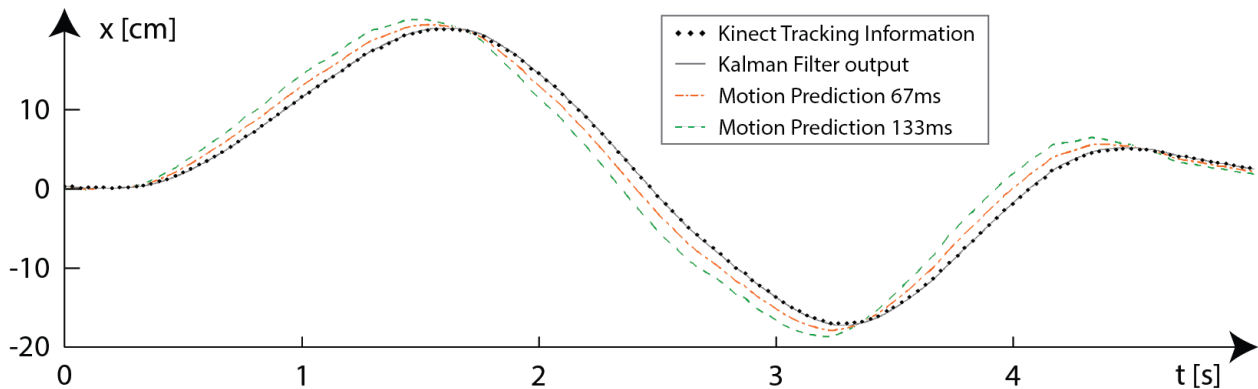


Figure 3: This plot compares predicted with measured face positions of a viewer. The dotted line represents the x-coordinates captured with full latency, where the dashed lines represent x-coordinates at different prediction time. Ideally, all three curves are identical with an offset on the temporal axis t , corresponding to the time predicted into the future. As clearly visible, the curves degenerate with increasing prediction time.

7. CONCLUSION AND FUTURE WORK

We have built a physical prototype to verify the conclusion of our comparison and to assess the quality of viewer tracking for motion parallax. Our display is able to show virtual persons behind the screen as if they were physically present. Immersion and realism are further increased by improving a Kinect face tracker with a Kalman filter and prediction. We are able to provide nearly correct motion parallax and correct perspective cues at high frame-rates with low latency.

To provide better immersion, a glasses-free flat-screen version of our display would be required. Furthermore, a hardware implemented eye tracker with higher refresh rate and lower latency could allow for faster viewer motion. Proof of concept is given with our prototype and we believe that transparent 3D screens will become prominent products on the display market in near future.

ACKNOWLEDGMENTS

This research, which is carried out at BeingThere Centre, is supported by the Singapore National Research Foundation under its International Research Centre @ Singapore Funding Initiative and administered by the IDM Programme Office.

REFERENCES

- [1] <http://www.google.com/glass/start/what-it-does/> (accessed 01.01.2014)
- [2] <http://www.samsung.com/us/business/displays/digital-signage/LH22NLBVLVC/ZA> (accessed 01.01.2014)
- [3] <http://www.lg.com/uk/commercial-display/lg-47TS30MF> (accessed 01.01.2014)
- [4] Sun, T., Wu, S. and Cheng, B., "Novel Transparent Emissive Display on Optic-Clear Phosphor Screen", SID Int. Symp. Digest Tech. Papers 44(1), 755-758 (2013).
- [5] Lee, B. and Hong, J., "Transparent 3D display for augmented reality", Holography, Diffractive Optics, and Applications V, (2012).
- [6] Mock, P., Schilling, A., Strasser, W. and Edelmann, J., "Direct 3D-collaboration with Face2Face - implementation details and application concepts", 3DTV-CON, (2012)
- [7] Kuster, C., Ranieri, N., Agustina, Zimmer, H., Bazin, J.C., Sun, C., Popa, T. and Gross, M., "Towards Next Generation 3D Teleconferencing Systems", 3DTV-CON, (2012).
- [8] Azarbajejani, A., Starner, T., Horowitz, B. and Pentland, A., "Visually Controlled Graphics", IEEE Transactions on Pattern Analysis and Machine Intelligence 15, 602-605 (1993).
- [9] Zhou, F., Duh, H., and Billinghurst, M., "Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR", Proceedings of the 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, 193-202 (2008).
- [10] <http://www.visionoptics.de/index.php?id=18&L=1> (accessed 01.01.2014)
- [11] <http://www.holopro.de/en/products/holoprotm.html> (accessed 01.01.2014)
- [12] <http://www.screen-tech.eu/en/ST-Professional-Trans.html> (accessed 01.01.2014)
- [13] Livingston, M. A., Sebastian, J., Zhuming, A. and Decker, J. W., "Performance measurements for the Microsoft Kinect skeleton", Virtual Reality Short Papers and Posters, 119-120 (2012).
- [14] Welch, G., and Bishop, G., "An Introduction to the Kalman Filter", <http://www.cs.unc.edu/~welch/kalman/Levy1997/index.html> (accessed 01.01.2014)