Toward the Light Field Display: Autostereoscopic Rendering via a Cluster of Projectors

Ruigang Yang[†], Xinyu Huang, Sifang Li, Christopher Jaynes

Center for Visualization and Virtual Environment, University of Kentucky, USA

Abstract

Ultimately, a display device should be capable of reproducing the visual effects that are produced by reality. In this paper we introduce an autostereoscopic display that uses a scalable array of digital light projectors and a projection screen augmented with microlenses to simulate a light field *for a given three-dimensional scene. Physical objects emit or reflect light in all directions to create a light field that can be approximated by the light field display. The display can simultaneously provide many viewers from different viewpoints a stereoscopic effect without head-tracking or special mechanical devices. We present a solution to automatically calibrate the light field display and an efficient algorithm to render the special multi-view images it requires by exploiting their spatial coherence. The effectiveness of our approach is demonstrated with a four-projector prototype that can display dynamic imagery with full parallax.*

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Virtual reality I.3.3 [Computer Graphics]: Display algorithms I.4.m [Image Processing and Computer Vision]: Projector calibration

1. Introduction

During the past several years we have seen tremendous progress in autostereoscopic displays: true three-dimensional display without the need for specialized and cumbersome gears is becoming feasible (see [Dod05] for a recent review). In addition to the commercialization of lenticular-based displays (e.g. [Ste03, X3D04]), 3D displays composed of multiple projectors show the greatest promise of delivering high-resolution 3D imagery in the near future. However, existing multi-projector 3D displays either only provide limited 3D effects (e.g. only horizontal parallax in [MP04]) or require special calibration equipment (e.g., a camera on a translating stage in [LIH*02]) that is difficult to set up or scale.

In this paper we present an autostereoscopic architecture and prototype that overcome the limitations of previous approaches. It is composed of an array of digital light projectors and a projection screen that is augmented with a sheet of microlenses. Projectors are used to generate an array of pixels at controlled intensity and color onto the screen and its array of microlenses. Each microlens (or lenslet) then transmits different colored light rays into different directions into a viewing volume in front of the screen.

Our proposed display in fact simulates an appropriate *light field* [LH96], therefore it is named the *light field display*. It can simultaneously provide many viewers from different viewpoints stereoscopic effect without head-tracking, head-worn lenses, or special gears. Figure 1 shows a prototype and its view-dependent effect.

While our light field display operates on the same principle as integral photograph [Ive31], its realization poses several unique and significant challenges. Unlike commercial 3D displays that rely on precise manufacturing, we seek to use computer vision techniques to automatically register various components in a global reference frame in which a light field can be generated. This approach brings significant advantages in terms of flexibility and scalability. Furthermore, a stereoscopic display is only interesting if it can display dynamic and interactive images. Our light field display requires

---

[†] This work is supported in part by US Department of Homeland Security and US National Science Foundation grant IIS-0448185.

Figure 1: *(Left) A prototype light field display composed of four projectors. (Right) View-dependent effects from the prototype. These images are captured with a regular camera from different viewpoints. The scene contains a Coke can, a teapot, and a textured background. The Coke can is rotated in the second row of images.*

images from many different viewpoints to be synthesized at once. We developed a rendering algorithm that takes advantages of the spatial coherence in a light field to dramatically reduce the rendering time and can help us achieve the scalability in the resolution required.

In summary we have made the following contributions to advance the state of the art in autostereoscopic displays:

- a complete calibration procedure to automatically register all the display components (i.e., projectors, the screen and microlens sheets) into a single global coordinate. Experimental results show that it is very accurate ($< 0.1\%$ relative error) and fast ($<$ 15 minutes for our four-projector setup).
- an efficient rendering algorithm to synthesize the special multiple-center-of-projection images for the light-field display. It reduces rendering time up to three orders of magnitude. In addition, it is simple to implement and can be fully accelerated with commodity graphics hardware.
- a light field display prototype that is capable of producing full color, full parallax, and full motion solid stereoscopic imagery, without resorting to headsets or user-tracking.

2. Display Calibration

In order to achieve an autostereoscopic effect, the light ray formation process should be controlled with great accuracy. This means the mapping between the pixels in the frame buffer and viewing directions in the three-dimensional viewing volume must be known. While commercial lenticular displays rely on precision manufacturing, we use computer vision techniques to discover this mapping information. This automatic procedure is critical to realizing a light field display.

As shown in Figure 2, a method to register different components (projectors, screens, microlens sheet etc.) in a single global coordinate frame is required. The calibration process can be divided into two parts. First, we need to establish *point correspondences* between each projector and the screen's reference frame. This can be accomplished by using existing camera-based multi-projector calibration techniques and we adopted the method from [RWC*98].

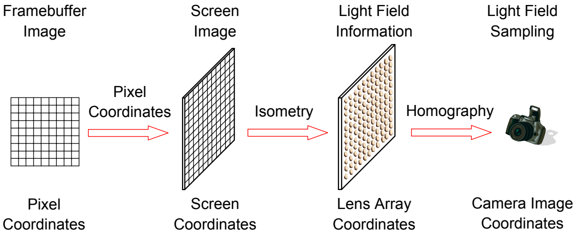

Figure 2: *The optic path in a light field display consists of the following components: the projectors' frame buffer, the diffuse screen, and the microlens sheet. A camera is used to observe the projected images.*

Secondly, we need to discover the geometric relationship between the screen and the lenslets on the microlens sheet. Each lenslet is modeled as a pinhole camera. We assume that all the lenslets have identical intrinsic parameters and their the centers of projections are on a plane, to which the optical axes are orthogonal. Therefore we only need to estimate a 2D translation of the optical center (or the image center on the focal plane). This is a necessary step for both regular lenslet layout (one offset for the entire screen) and irregular lenslet layout (one offset for each lenslet).

To estimate the translational offset for each lenslet, we exploit a simple optic property: when a group of rays parallel to the optic axis hit a convex lens, they will focus on the image center on the focal plane. So if the screen is illuminated with a directional light source, we can use a camera to observe the focus point (i.e., a bright dot) to estimate the 2D offset of a lenslet. This method is simple and efficient: the entire microlens screen can be calibrated with just a single image.

In summary, we place a camera behind the screen (the diffuse side) to measure the unknown calibration parameters for a given setup. Using fiducial points on the screen we estimate the homography between the camera image plane and the screen. Then we find the projector-to-screen mapping using structured light and the 2D offsets for all the lenslets using a directional light source. All the calibration results are defined in the metric screen coordinate system formed by the fiducials.

3. Multi-View Rendering

The imagery behind the microlens screen is a composite of many images of the same scene from different viewpoints, i.e., a Multiple-Center-of-Projection (MCOP) image [RB98]. The corresponding pixels behind each lenslet, collectively referred to as a *screenlet*, form a tiny perspective view of the scene. Forming a MCOP image from a light field is an efficient process that has been successfully applied to real-time rendering of dynamic scenes (e.g. [MP04]). We focus our research on efficient rendering of MCOP images from a computer-graphics model.

We present an algorithm that exploits the regular layout of viewpoints. Given the regular layout in the microlens array in the layout, there are many parallel view rays among these screenlets. A view ray, denoted as $P^k_{(x,y)}$, is defined by a pixel at (x,y) in the screenlet C_k and the center of projection of C_k. Note that (x,y) is defined in the local image coordinate of a screenlet. For a fixed (x,y), $\{P^0_{(x,y)}, P^1_{(x,y)}, \cdots, P^n_{(x,y)}\}$ are all mutually parallel. So instead of rendering each perspective screenlet, we can render each group of parallel rays using orthographic projection (*Stage I*) and then reassemble the MCOP image (*Stage II*), as shown in Figure 3. We call this technique *Parallel-Group Rendering* (PGR).

PGR requires m rendering passes where m is the resolution per screenlet while traditional single-view perspective rendering (SVR) requires n rendering passes where n is the total number of screenlets (or lenslets). Since in a typical 3D display the total number of screenlets is a few orders higher than the resolution of each screenlet, PGR brings tremendous savings in rendering time. While this idea may seem simple, to the best of our knowledge, it has not been introduced in the context of light field generation. In addition, we have developed techniques to ensure the accuracy of the final image.

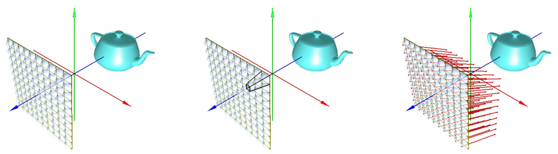

Figure 3: *Parallel rays in a microlens array.* ***(Left)*** *A teapot scene is to be visualized.* ***(Middle)*** *A traditional way to render MCOP images is to render the perspective view for each screenlet one by one.* ***(Right)*** *Our novel rendering algorithm renders groups of parallel rays, significantly reducing the number of rendering passes needed.*

4. Results

We have built a prototype system with four projectors (shown in Figure 1 left). All projectors are connected to a PC with two NVIDIA GeForce7800 graphics cards (each driving two). Our screen is composed of four 8 × 10 inch microlens sheets. A close-up lens is attached to each projector so it can be focused with a very short throw-distance. In our prototype, the screen size is 241 × 178 mm with 103 × 89 lenslets. Each lenslet's diameter is 2.286 mm with a field of view of 45 degrees.

Calibration We use a six-megapixel digital camera (Nikon D70) to perform calibration. The radial distortion in camera is removed a priori to provide calibration using our linear technique. Although our screen is planar, nonlinear projector distortion cannot be ignored, especially with the insertion of the close-up lens. Therefore the mapping from projector to screen is described as a look-up function that encodes both the projective transformation and non-linear projection distortion. For lenslet calibration, the light source is over four meters away. Based on the calibration results, we estimated that the mean diameter for each lenslet is 2.2886 mm, which is within 0.1% of the specification (2.286 mm).

Rendering We use the PGR algorithm to drive the light field display. Some of the final results are shown in Figure 1. The view-dependent effect is obvious. **Video results can be found at** `http://www.vis.uky.edu/~xhuan4/LightFieldDisplay.html`. The animation is from pre-computed images.

Figure 4 demonstrates our system's capability to produce "out-of-screen" stereoscopic effect. The virtual scene consists of two balls: the red ball is 25 mm in front of the screen, and the green ball is behind the screen. In our test we set up a real plummet in front of the screen onto the virtual red ball. The plummet and the red ball are registered together when we move the viewpoint.

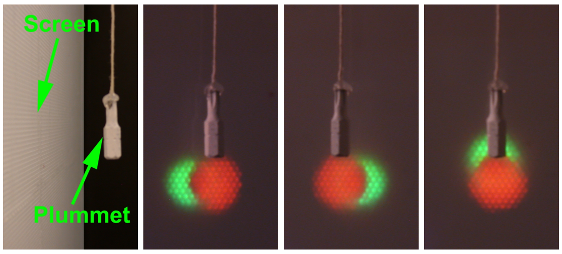

Figure 4: *Demonstration of "out-of-screen" effect. The left-most image shows a side view of the setup, in which a real plummet is in front of the screen. The virtual red ball appears to be registered with the plummet.*

4.1. PGR's Accuracy and Performance

We further demonstrate the accuracy and speed of our PGR algorithm. We for now assume the lenslets are rectangular and on a regular grid. We first rendered 16 views (on a 4 × 4 grid) at 256 × 256 resolution using a car model with 60,698 triangles. The image differences of PGR and SVR are shown in Figure 5. The final images from PGR are almost identical to these rendered with SVR.

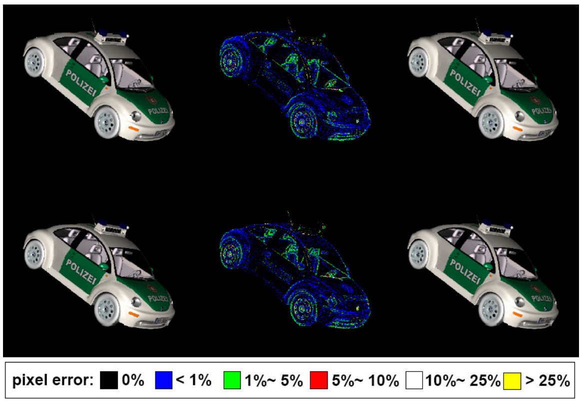

Figure 5: *The two images on the first column are rendered by SVR. The third column is rendered by PGR. The second column is the difference between the first and third column.*

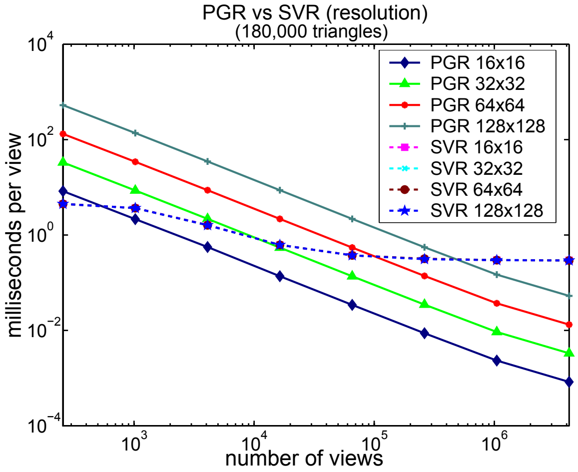

Figure 6: *Comparison between PGR and SVR for different resolution of lenslet using 180,000 triangles. The speed curves from SVR are identical for different resolutions because the image size is rather small.*

In terms of speed, the timing results of PGR and SVR are summarized in Figure 6. The tests were performed on a Intel 3.2GHz PC with 1GB RAM and NVIDIA GeForce 6800 card. In Figure 6, PGR's rendering time per view is monotonically decreasing as that of SVR quickly becomes constant. The break-even point depends on the per view resolution, which equals to the number of rendering passes for PGR. We can see that even with a high-resolution screenlet, PGR can outperform SVR when the number of views is over 0.5 million. That is the number of "pixels" a viewer can see at any given viewpoint. Today's 2D display can easily achieve over 1 million pixels. Therefore PGR's advantage is particularly significant for high-resolution displays.

Most related to PGR is the multi-view rendering (MVR) algorithm by Halle [Hal98]. It was reported that MVR can achieve one or two orders of acceleration for large number of views, which is similar to that from PGR for low to medium resolution screenlets. From an algorithmic standpoint, PGR is more flexible in handling of irregular grid. In addition, PGR is much easier to implement.

5. Conclusion

We present a novel architecture to create autostereoscopic displays from commodity components. By using computer-vision algorithms to automatically align a multi-projector display with off-the-shelf microlens screens, our approach provides an unprecedented level of flexibility. The size, resolution, and viewing volume of our light field display can be easily changed to satisfy the end user's requirements and specification. In addition, we have developed a novel rendering algorithm that exploits the special structure of the rendering task as well as the acceleration capabilities provided by commodity graphics hardware. It increases the rendering speed by one to three orders of magnitude.

Looking into the future there are many exciting places for improvement. One immediate step is increase the lenslet resolutions to present a more compelling 3D display. We have not addressed the photometric issues in this paper. This may become problematic as the lenslet resolution increases. Furthermore, we are in the process of building a programmable pixel routing hardware so that a cluster of PCs can distributively render Stage-I images without the extra cost of composition. This will lead to a truly *interactive* 3D display that can provide full motion parallax.

References

[Dod05] DODGSON N. A.: "autostereoscopic 3d displays". *Computer 38*, 8 (2005), 31–36.

[Hal98] HALLE M. W.: Multiple Viewpoint Rendering. In *Proceedings of SIGGRAPH 1998* (1998), pp. 243–254.

[Ive31] IVES H. E.: Optical Properties of a Lippmann Lenticulated Sheet. *Journal of Optical Society of American 21* (1931), 171–176.

[LH96] LEVOY M., HANRAHAN P.: Light Field Rendering. In *Proceedings of SIGGRAPH 1996* (New Orleans, August 1996), pp. 31–42.

[LIH*02] LIAO H., IWAHARA M., HATA N., SAKUMA I., DOHI T., KOIKE T., MOMOI Y., MINAKAWA T., YAMASAKI M., TAJIMA F., TAKEDA H.: High-Resolution Integral Videography Autostereoscopic Display Using Multi-Projector. In *Proceedings of the Ninth International Display Workshop* (2002), pp. 1229–1232.

[MP04] MATUSIK W., PFISTER H.: 3D TV: a Scalable System for Real-time Acquisition, Transmission, and Autostereoscopic Display of Dynamic Scenes. *ACM Trans. Graph. (SIGGRAPH Proceeding) 23*, 3 (2004), 814–824.

[RB98] RADEMACHER P., BISHOP G.: Multiple Center-of-Projection Images. In *Proceedings of SIGGRAPH 1998* (1998), pp. 199–206.

[RWC*98] RASKAR R., WELCH G., CUTTS M., LAKE A., STESIN L., FUCHS H.: The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. *Computer Graphics 32* (1998), 179–188.

[Ste03] STEREOGRAPHICS CORPORATION: SynthaGram Monitor, 2003. http://www.stereographics.com.

[X3D04] X3D TECHNOLOGIES CORP: X3D Displays, 2004. http://www.x3d.com/.