SOME PRACTICAL ASPECTS OF LIGHT FIELD DISPLAY SYSTEMS

Phil SURMAN(1), XIA Xinxing(1) , SONG Weitao(1), ZHANG Xiangyu(1), WANG Shizheng(2)

(1) Nanyang Technological University
(2) Chinese Academy of Sciences

ABSTRACT

Any developments in light field imaging and processing are obviously only useful if there are light field display technologies available on which to show it. In the Advanced Displays Laboratory (ADL) at Nanyang Technological University (NTU), several embodiments of these are under development. The multi-layer Tensor Display is a direct-view display based on multiple transparent liquid crystal display layers that produce image depth at continuously variable planes. The super multiview variants are multiview displays with high view density that effectively produce light fields at the users' eyes. The near-eye augmented reality display with focus cues solves the serious accommodation/convergence problem which is the principal cause of discomfort with this type of display. Finally, an interesting new development is that of vision correction at the display without the user having to wear spectacles. In this paper, we consider some of the practical aspects of bringing these developments to completion.

Index Terms — Light field display, Tensor Display, super multiview display, augmented reality (AR), virtual reality (VR) motion parallax, accommodation/convergence conflict

1. INTRODUCTION

First, it is useful to consider the reasons for needing a light field display. Three principal reasons are; the presentation of stereoscopic images, elimination of accommodation/ convergence (AC) conflict [1] and to display motion parallax (the ability to 'look around' objects). Depth can be perceived in images using just stereo disparity, which is two slightly different perspectives for each eye. Disparity is the strongest stereoscopic cue and in many usage cases it is sufficient.

In the real world, different parts of the perceived image are at different distances from the viewer but this does not create a problem as the convergence of the eyes naturally matches the distance at which the eyes focus. However, when a 3D image is shown on a screen, the eyes converge at the apparent distance of an object, which is set by the disparity, but they still focus on the screen, as shown in Figure 1. There are various criteria for the limit of comfortable viewing, one is to keep the difference between focus and accommodation to within 0.3 diopters [2] and another being the 1° difference rule [3]; in fact, these two are virtually the same. The effect of AC conflict is sometimes overstated but certainly some disparity limitation is required.

Motion parallax is not a strong stereoscopic cue and in many cases it is unnecessary; for example, it is of little use for a seated audience. On the other hand, if the audience is in a collaborative environment, for example in a virtual surgical operation, then it would be very useful if the image occupies the same apparent volume in space for every user. The position of users hand or surgical instrument inserted into the virtual image would appear to be the same in relation to the virtual object to all users. In this case, a light field display is the only suitable type for a compact display. In principle, it is possible to use a multi-user head tracker and a fast display, and in fact this has been considered by the authors for a collaborative table-top display. However, for N users, 2N times the normal frame rate is required, which precludes the use of any current flat panel display.

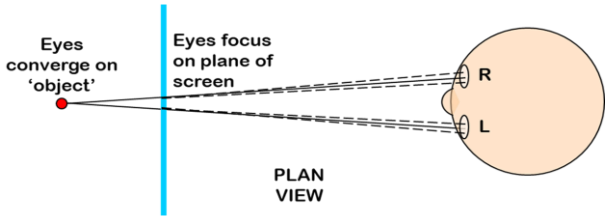
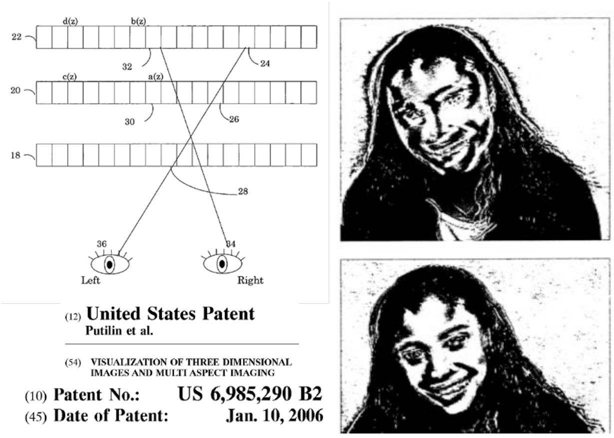


Figure 1. Accommodation/convergence con-

2. BRIEF OVERVIEW OF PREVIOUS WORK

Light field displays replicate the light pattern of the original scene using geometrical optics principles rather than those of light interference, such as holography that are more complex to implement. When a scene is observed from inside a convex hull, every light ray passing through the hull can be specified by what is known as a four-dimensional plenoptic function [4]. In displays, the observation region is not the complete hull but the plane of the screen that is bounded by its sides. The simplest way of specifying any particular ray is the X, Y position of the ray exiting the screen and the horizontal and vertical components of its angle. The term "light field" was first coined by Andrey Gershun in 1936 [5].

Before MIT brought multi-layer light field research to prominence, it had originally been developed in Russia in the early 2000's [6] and used neural networks to generate the images on a stack of transmitting panels. Fig. 2 is from a 2006 patent [7] and a product based on this, known as SmartrON was marketed by Neurok Optics where two liquid crystal panels and a special mask were arranged in a stack.



(12) **United States Patent**
Putilin et al.

(54) VISUALIZATION OF THREE DIMENSIONAL IMAGES AND MULTI ASPECT IMAGING

(10) **Patent No.: US 6,985,290 B2**
(45) **Date of Patent: Jan. 10, 2006**

Figure 2. SmartrON multi-layer display patent drawings

The multilayer display approach was then taken up by the MIT Media Lab where there are various embodiments including tensor displays [8], high rank display (HR3D) [9] and dynamic parallax barrier displays [10]. The high-rank 3D display uses a stack of semi-transparent LCD layers and is able to generate different light rays in each direction from the same point of the screen. Each LCD layer contains a different pattern and the whole stack can approximate the light field of a given scene and re-create the scene with a hologram-like appearance.

The patterns on each LCD can be obtained through a computationally expensive optimization as the light field of a natural 3D scene can be described by a set of low-rank matrices.

The Holografika display [11] uses optical modules to provide multiple beams that converge and intersect in front of the screen to form real image “voxels”, or diverge to produce virtual voxels behind the screen. A similar approach has been taken by the University of Southern California; they have built a 110° field of view system using 72 Texas Instruments DLP pico projectors [12].

In super multiview (SMV) displays, a dense series of perspective views is produced across the viewing field. The aim of several research teams is to have a sufficiently dense set of views in order to enable the users’ eyes to accommodate and so avoid AC conflict. Sun Yat Sen University has developed a system with OLED microdisplays [13] that produces viewing zones 1.64 mm wide. ETRI in South Korea has determined, from a setup with multiple projectors and using a single subject that a 2.05 diopters accommodation range can be obtained from providing four views per eye pupil.

3. LIGHT FIELD DISPLAY DEVELOPMENT AT ADL

Four types of light field display have been developed at ADL. The Tensor Display is a direct-view multi-layer display based on MIT’s hardware configuration. Three SMV demonstrators have built that show the operation of the specially developed beam steering screen and fast temporal multiplexing display. An AR display has been developed where the images from a 2-layer light field display are overlaid over the real world. Work is also underway on glasses-free vision correction where localized light fields are formed around the eyes and that compensate for eye aberrations.

3.1 Tensor Display

The display hardware shown in Figure 3 comprises three liquid crystal display (LCD) panels removed from their monitor housings and the standard backlight replaced by a matrix of 1,500 high-power white LEDs. All the display panel polarizers, with the exception of the one closest to the illumination source, are removed and the nearest polarizer to the viewer is replaced with a non-diffusing type.

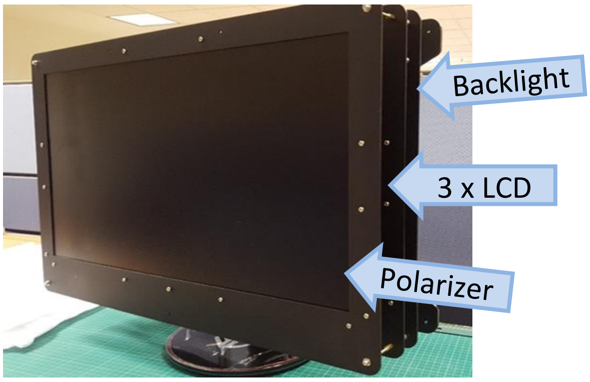


Figure 3 Tensor Display

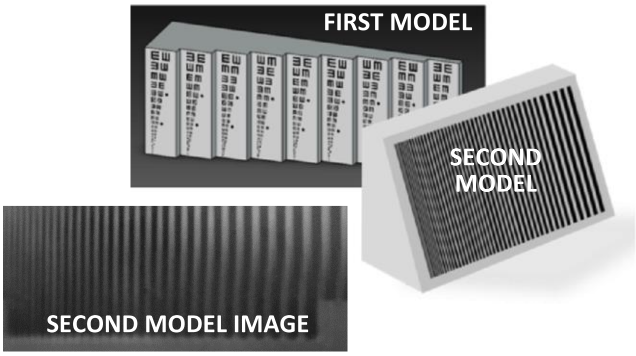


Figure 4 Virtual models used for Tensor Display characterization

Although the hardware design is similar to MIT’s version, the Tensor Display proved to be a very useful testbed for developing light field display system characterization techniques where the complete capture, processing and display components.

Initial work was carried out using 12 subjects who registered results on score sheets. This was a lengthy process and the results were not as accurate as they might have been, principally due to the virtual object displayed that was produced in Solidworks being unsuitable (Figure 4 *top left*). An improved model (Figure 4 *top right*) was made for the second iteration where the sloping front surface gives continuous depth and randomly orientated Snellen E’s are replaced by a bar grating whose pitch increases by a factor of $2^{1/32}$ between black and white bar centers (16 cycles/octave).

The resulting image pattern is shown in Figure 4 *lower* and close examination reveals that it approximates a sinusoidal grating pattern. This, along with the presence of ringing artefacts on the images of the Snellen Es indicates that the overall system can be considered as being a low-pass filter which has the advantages that the spatial resolutions can be expressed in terms of cut-off frequencies and these can be determined by measuring contrast ratios on a photograph of the image of the model. The procedure is detailed in Reference [15]. A summary of the results of this paper is as follows:

• Spatial cut-off frequency	554.4	cycles/m
• Angular cut-off frequency	8.8	cycles/radian
• Depth of field (DOF)	17.5	mm
• Contrast ratio	14.6:1	

All the parameter values are fairly poor. Although it is not immediately obvious why the spatial frequency, depth of field (DOF) and contrast ratio values are low, the low angular resolution can principally be attributed to the sparse nature of the 7 x 7 matrix of input images produced from a virtual camera array at 1 m distance with a spacing of 50 mm.

3.2 SMV displays

SMV displays only show motion parallax in the horizontal direction and are therefore referred to as horizontal parallax only (HPO). As mentioned in Section 2, when the discrete views are sufficiently dense, accommodation is provided. The need for such a high density is arguably not necessary at typical cinema or television viewing distances as the difference in near and far displayed distances in terms of diopters will generally be within the range of comfort. Also, with today’s technology, it would not be possible.

Therefore, the realistic goal of an SMV display is not to provide accommodation cues but to give a good depth of field in the image by having a high horizontal angular resolution. The need for this

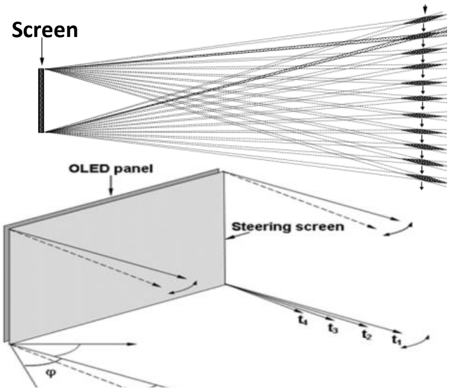


Figure 5 SMV display using steering screen

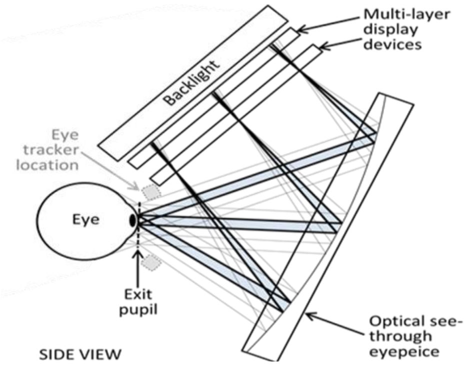


Figure 7 Light field AR schematic diagram

is made apparent in multiview displays where images become blurred for objects appearing away from the screen, with multiple edges being observed.

A free-viewing 3D display of any type must possess angular resolution where the emitted light from a point on the screen varies with angle. In most cases, apart from examples such as integral imaging and holography, the light only varies in the horizontal direction.

A free-viewing 3D display of any type must possess angular resolution where the emitted light from a point on the screen varies with angle. In most cases, apart from examples such as integral imaging and holography, the light only varies in the horizontal direction.

Where only spatial multiplexing is used, angular resolution is obtained at the expense of spatial resolution. This effect is very apparent in multiview displays. Consider the example of a 9-view, 1920 x 1080 slanted lenticular display. The actual figures will depend on the sub-pixel mapping but let us assume that the horizontal loss and the vertical resolution loss are both a factor of 3 (3x3 = 9: the number of views), each perceived image is only 640 x 360.

As this is clearly unacceptable and is the reason for slanted lenticular technology never being widely adopted, ADL decided to develop spatio-temporal multiplexing where a high-speed display could be used instead of sheer brute-force numbers of pixels (for example, Reference [12]).

This is achieved by producing a series of sparse viewing zones with voids in between them (Figure 5 *top*) and filling in the voids over time using a light steering screen located directly in front of the display screen (Figure 5 *lower*).

The key components are the steering screen and a fast display. The steering screen, comprising a prism array used as one of the substrates of a birefringent liquid crystal layer, was developed at ADL. Light is steered through two angles by rotating the input polarization using a pi cell. More angles are obtained by cascading the screens. The fast display is a 2.8” OLED panel obtained from Visionox (Figure 6 *left*). This has connections directly on to the glass, thus bypassing the speed bottleneck of the display driver chip. This runs at a frame rate of 1 kHz using an in-house designed digital display driver.

As the fastest direct-view panel we could obtain is small, another version was built using fast digital light processor (DLP) pico projectors (Figure 6 *right*). 12 projectors are used and two cascaded steering screens are used to increase to effectively 36 projectors using 3 x temporal multiplexing. 24 virtual projectors are produced between the 12 real projectors.

3.3 Light field AR

Currently, near-eye display (NED) performance is generally unsatisfactory; resolution tends to be low, particularly in displays using a single panel, such as those where smartphones provide the images. The appearance of a visible pixellation structure is referred to as the “screen door” effect. A perceived resolution at the eye of 30 cycles per degree is desirable as this corresponds to 20/20 vision; conveniently, it is also equivalent to the image of a single pixel subtending an angle of one arc minute at the eye.

For a truly immersive experience, a wide field of view (FOV) is important for giving the sensation of being immersed within a scene, rather than just observing it through a window. Ideally, the complete peripheral vision region is filled so that it extends to around 200° horizontally and 110° vertically. In AR, where the virtual world is superimposed over the real world, the left and right peripheral zones can be substantially provided by the real world.

The need for accommodation is even greater for AR than for VR as the virtual image position must match that of the real world that it overlays. This is typically from 250 mm to infinity; a range of 4 diopters. This is a severe problem and consequently the NED research of ADL is concentrated in this area.

We use a two-layer version of the Tensor Display where in this case, the images are optimized to give a large DOF that is visible over the small area, which is the exit pupil of the optical system.

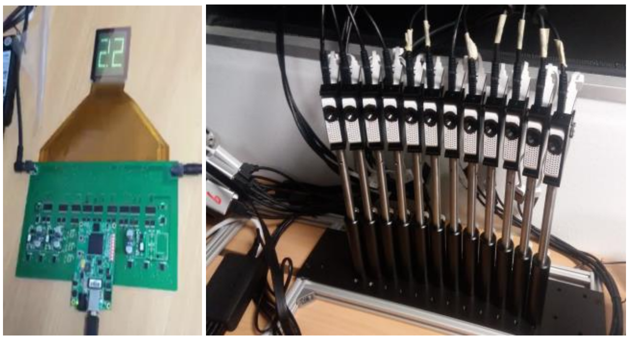


Figure 6 SMV prototypes: *Left*-1 kHz OLED, *Right*-36-view

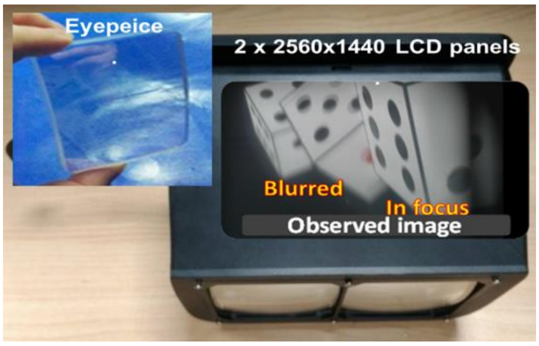


Figure 8 SMV Light field AR showing effect of focusing

The real world is seen through a half-mirror combiner that also serves to focus the image from the two 2560 x 1440 LCD screens as shown in Figure 7. A prototype has been built (Figure 8) that gives 2 diopters focusing range.

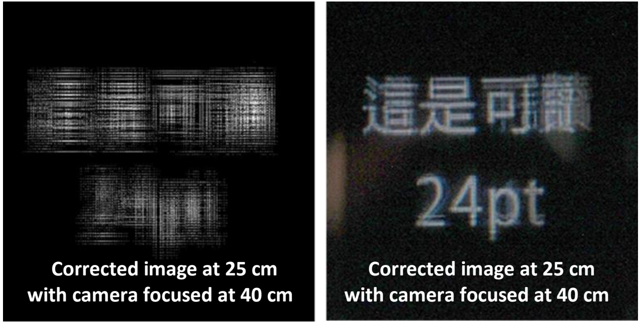


Figure 9 Glasses-free vision correction

3.4 Glasses-free vision correction

The concept of vision correction without the use of spectacles has been around for several years [16] but until now, it has not got to the stage of being commercialized.

Conventional displays control the intensity of every pixel to create 2D images. Light from each pixel emits substantially the same illumination over large angles and each pixel is effectively a tiny light source. Our eyes automatically focus on the screen in order to form a sharp image on the retina. By adding an array of microlenses on top of a screen, we can also control the direction of the light emitted.

Leveraging on these current advancements in the field 3D and the increasing resolution of the displays, we will correct vision problems digitally from the screen itself and users with myopia, hyperopia, or presbyopia do not need to wear correction glasses. By placing the pixels at the microlens focal plane, the emitted light from a pixel becomes a parallel beam.

The position of the pixel with respect to the lens determines the direction of the beam. The size of the micro-lens in the square array determines the number of light rays controlled; if each lens covers 5 x 5 pixels then we have 25 light rays for each lens.

There is a trade-off between correction quality and the resolution of the corrected image. However, the recent evolution of display technology is creating higher screen resolutions that exceed the requirement of human eyes at the normal reading distance of around 25 cm.

With the hardware in place, a well-defined light direction for any individual pixel is pre-determined and by having the user's vision prescription from an ophthalmologist, we can calculate exactly where these light beams go to on the user's retina. Summation of all these beams on retina forms an image that our eye sees. Mathematically, we build an exact map from screen's image to the image on retina through a linear transform T that is very well defined by three factors; microlens hardware, the screen pixel configuration and the conditions of the user's eyes. All three factors are known to our developers from the lens design, smart phone manufacturer and ophthalmologist, so it can be easily established.

Calculating the necessary inverse transform is difficult for a very large matrix and various optimization processes are implemented to do the approximation. However, the calculation is one-time process, and the result is stored for lifetime use.

Initial work has been carried out with pinhole arrays in the place of microlenses. Figure 9 *left* shows an image on a screen at 25 cm that has been processed by the algorithms developed and Figure 9 *right* shows the same image seen through a prototype pinhole array with the camera focused at 40 cm; this represents a 1.5 diopter correction.

4. CONCLUSIONS

A summary of the findings from our research is as follows. The Tensor Display would benefit enormously from having a denser set of input views, or even better still, input with continuous parallax. The SMV displays developed do not have sufficient viewing zone density for focus cues and, at present, there is no suitable fast direct-view display panel available. However, micro LEDS show promise for this application provided they have an appropriate display driver. The steering screen is problematic as it is difficult to scale-up size and also, to run at a sufficiently high frame rate. The light field AR gives a good focusing range of 2 diopters that will be improved when the number of LCD layers is increased to three or more. Vision correction research is in its early stages, but we are optimistic we can solve it given our experience with other light field display types.

5. REFERENCES

[1] D. M. Hoffmann, A. Girshik, K. Akeley et al. "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," J Vis, Vol. 8 No. 3 Article 33, 2008

[2] F. W. Campbell, "The depth of field of the human eye," Journal of Modern Optics, 4, Vol. 4, pp167-164, 1957

[3] W. IJsselsteijn, et al. "Perceived depth and the feeling of presence in 3DTV," Displ. 18, Elsevier, pp 207–214, 1998.

[4] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz & P. Hanrahan, "Light Field Photography with a Hand-held Plenoptic Camera," Stanford Tech Report CTSR 2005-02 http://hci.stanford.edu/cstr/reports/2005-02.pdf

[5] A. Gershun, "The Light Field," Journal of Mathematics and Physics, Vol, 18, Issue 1– 4, John Wiley & Sons, Ltd, 1939

[6] A. Loukianitsa and A. N. Putilin, "Stereodisplay with neural network image processing", Proc. SPIE 4660, Stereoscopic Displays and Virtual Reality Systems IX, 207 (May 2002),

[7] US 6985290 B2, "Visualization of three dimensional images and multi aspect imaging"

[8] G. Wetzstein, D. Lanman, M. Hirsch & R. Raskar, "Tensor Displays: Compressive Light Field Synthesis using Multilayer Displays with Directional Backlighting," ACM Transactions on Graphic, Vol. 31 Issue 4, July 2012

[9] D. Lanman, M. Hirsch, Y. Kim & R. Raskar, "High-Rank 3D Display using Content-Adaptive Parallax Barriers," http://alumni.media.mit.edu/~mhirsch/hr3d/

[10] D. Lanman, M. Hirsch, Y. Kim & R. Raskar, "Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization," ACM Transactions on Graphics, Vol. 29, No. 6, p. 163, 2010

[11] US Patent No. 6,201,565 B1 "Method and apparatus for displaying three-dimensional Images," T. Balogh, 2001

[12] K. Nagano, A. Jones, J. Liu, J. Busch, X. Yu, M. Bolas & P. Debevec, "An Autostereoscopic Projector Array Optimized for 3D Facial Display," Siggraph 2013 Emerg. Tech.

[13] L. Liu, Z, Pang and D. Teng, "Super multi-view three-dimensional display technique for portable devices," Optics Express 24(5):442, Mar. 2016

[14] B.-R. Lee, et al. "Increasing the depth of field in Multiview 3D images," Proc. SPIE 9867, Three-Dimensional Imaging, Visualization, and Display 2016, 98670T (Jun. 2016)

[15] P. Surman, S. Wang, J. Yuan & Y. Zheng, "One-stop measurement model for fast and accurate tensor display characterization," J Opt Soc Am A Opt Image Sci Vis. 2018 Feb 1;35(2):346-355

[16] F. Huang, G. Wetzstein, B. Barsky & R. Raskar, "Eyeglasses-free Display: Towards Correcting Visual Aberrations with Computational Light Field Displays," ACM Trans. Graph. (Proc. Siggraph), vol. 33, no. 4, pp. 1-12, Jul. 2014.