Three Dimensional Lenticular Display Rendering Based on Light Field Acquisition

Renjing Pei, Zheng Geng, Kui Ma, Zhang Mei, Wang Rong
Institute of Automation, Chinese Academy of Sciences,China.

Abstract
Crosstalk is a critical defect affecting quality in 3D displays. Existing methods require tedious computations or device-specific optical measurements and results are often sub-optimal for 3D productions. We propose a method based on light field acquisition and optimization for crosstalk reduction. Algorithms were developed and experimental results showed superior performance.

Author Keywords
Lenticular 3D display; Crosstalk optimization; Light field; Light field capture.

1. Introduction

Crosstalk is a critical defect affecting image quality in multiview lenticular 3D displays. The inability to completely isolate left/right image channels for left/right eyes often results in crosstalk - the content from one channel is partly presented in another channel [1]. Two different approaches are commonly used to reduce the crosstalk effects: (1) to incorporate extra elements, e.g., parallax barriers, pixel masks or add a timing control scheme; (2) to correct subpixel values in the synthetic images to minimize the crosstalk. Examples include crosstalk coefficient calculation [2] or weighted value [3]. All of these methods reduce crosstalk in the viewpoints reasonably well. However, drawbacks of these approaches include (1) device-specific: they need precise measurement of optical parameters for each display device in order to maximize the crosstalk reduction or (2) Tedious and error-prone computation. Hence, the efficiency display device establish and detective results are often sub-optimal for high-quality 3D display productions.

To solve this problem, we propose a light field concept to optimize lenticular 3D displays. The light field, proposed by Levoy and Hanrahan [4], characterizes the radiance flowing through all the points in each possible direction [5] The ultimate goal of 3D display systems is rendering a “perfect” display – reproducing the light field L generated by real world physical objects, as if the objects contained in 3D images are placed in front of the lenticular sheet. Experimental results validate our method is simple and superior.

2. Method

Overview of the proposed method: We developed algorithms for optimization based on light field framework. Experiment is conducted for multiview rendering- N-views[1] ($N \in \{t|t \in Z^+, i = t/9, j = \sqrt{t}, i \in Z^+, j \in Z^+\}$) on a 9-view 3D display hardware, with the resolution of $H \times W$ and field of view (FOV) of 15.3^{o}. In this paper, the continuous light field in the real word is sampled to a discrete light field. We use $II^{\sim}, II, L^{\sim}, L$ to represent the initial synthetic image, target synthetic image, emitted light field captured by camera and target light field, respectively. When synthetic image displayed on LCD plane, there are $H \times W \times 3$ rays emit from the LCD plane that pass through the lenticular sheet (in Fig 2.(a)). Correspondingly, $H/\sqrt{N} \times W/\sqrt{N} \times 3$ rays could be captured at the nth position. $H/\sqrt{N} \times W/\sqrt{N}$ represents the resolution of the point light field image $l_{mn}^{\sim}$ ($m = 1{\sim}N, n = 1{\sim}N$) photographically acquired by the camera after image pre-processing.

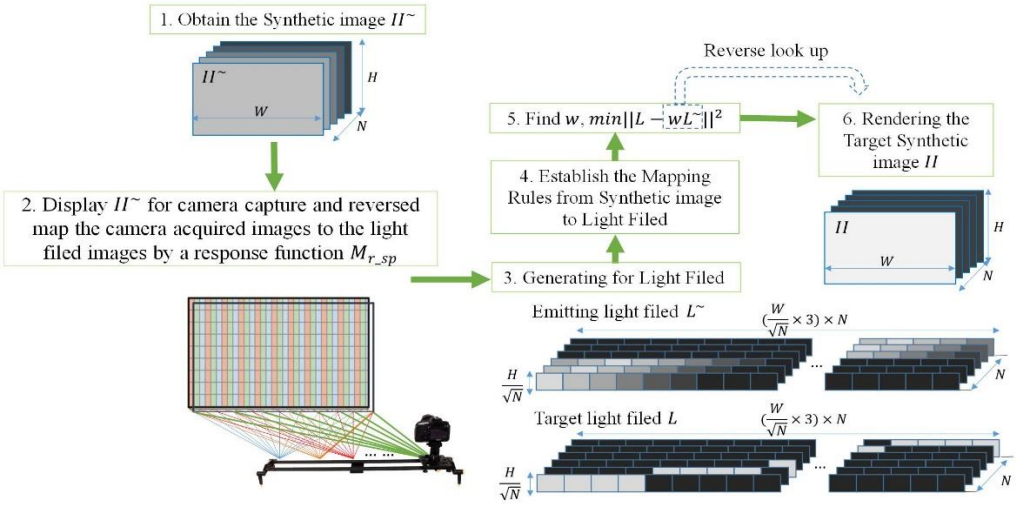

Figure 1. Overview of our proposed light field decomposition method.

[1] *In our experiments, N = 9 or 36. When N =36, it’s called a super multiview rendering*[2]

[2] *M-view display hardware for N-view multiview display ($M < N$), with M viewpoint views and (N-M) non-viewpoint views.*

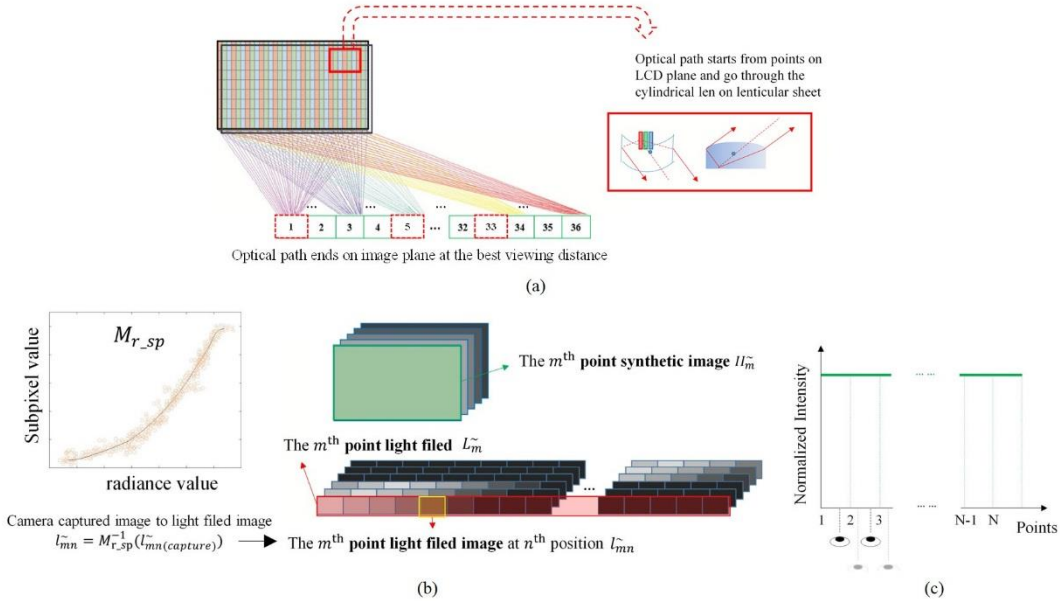

Figure 2. Overview of our proposed light field decomposition method. (a) Pre-set camera capture positions (N), including 9 pre-set viewpoints marked with red squares and $(N-9)$ intermediate non-viewpoints with green squares. We used $H \times W \times 3$ rays to approximate infinite number of rays in a real light field. These rays are emitted from the display device LCD and passed through the lenticular sheet to the image plane. $H/\sqrt{N} \times W/\sqrt{N} \times 3$ rays can be captured at the nth position ($n = 1 \sim N$). (b) $L_m{}^{\sim}$ in red square denotes the mth point light field, which contains N light field image $l_{mn}^{\sim}$; and $II_m{}^{\sim}$ in green square denotes the mth point synthetic image. (c) No residual crosstalk imagination intensity distribution map for N-view hardware display system.

Our propose is to produce the emitted light field $L^{\sim}$ and optimize it to be most similar to the target light field L. A schematic of the technique is shown in Fig 1. In our algorithm, there are six-stage procedures for the final synthetic image rendering:
(1) Obtain the initial synthetic image $II^{\sim}$: 1) synthetic image of each point $II_m{}^{\sim}(m = 1 \sim N)$ is rendered when the mth view image turned on at maximal intensity while others turned off (refer to Philips' method [6]; 2) we obtain the initial synthetic image aggregated from $II_m{}^{\sim}$ as: $II^{\sim} = \cup_{m=1}^{N} II_m{}^{\sim}$. $II^{\sim}$ contains information of N points synthetic images; (2) Capture for the light field: display each $II_m{}^{\sim}$ and photographically acquire images with a moving camera recorded as $l_{mn(capture)}^{\sim}$ after capture images' pre-processed (anti-distortion, 2D projection and interesting region extraction). A nonlinear mapping $V_{sub} = M_{r\_sp}(X)$ from radiance value X to acquired image subpixel' value V_{sub} for light field images $l_{mn}^{\sim}$ [7]; (3) Generate the light field: 1) all the light field images in (2) are combined for an emitting light field $L^{\sim}$; 2) the actual camera positions for each movement is exactly calibrated during the photographically acquisition in (2) and the target light field L can be computed according to the actual camera positions; (4) Establish the mapping rules from synthetic image displayed on LCD plane to the viewing plane: we determine s an one-to-one mapping (recorded in a look-up-tabel) from points on LCD plane (synthetic image) to viewing plane by rays back-projection method in [3]; (5) Compute the optimization matrix: find the optimization matrix w: $\min||L - wL^{\sim}||^2$. We solve the optimization problem by a Maximum Likelihood (ML) method – seeking a w that maximizes the likelihood function $P(L|w)$; (6) Render the target synthetic image II: with an optimal matrix w, the optimization synthetic image II can be reversely searched explicitly from $wL^{\sim}$. Details are shown in following sections.

Emitting and Target Light Field Acquisitions: We get the emitting light field by camera as follow: first, we determine 9 viewpoints of 9-view hardware in front of the display device, at the best viewing distance. And $N/9$ equal divisions are defined between two neighbor viewpoints, providing 9 viewpoints and $N-9$ non-viewpoints (in Fig. 2(a)). Then, each point synthetic image $II_m{}^{\sim}$ are displayed and photographically aquired at N positions in turn by a fixed focus moving camera. For example, when the camera at position 1, point synthetic image $II_m{}^{\sim}(m$ from 1 to $N)$ are displayed sequentially, and we acquired and recorded each light field image as $l_{m1}^{\sim}$(after preprocessing). The emitted light field $L^{\sim}$ is aggregated as $L^{\sim} = \cup_{m=1}^{N} L_m{}^{\sim} = \cup_{m=1}^{N} \cup_{n=1}^{N} l_{mn}^{\sim}$. $L_m{}^{\sim}$ contains information of all N-point light fields.

Since the pre-set positions are not practical during the camera movements, when we compute the target light field, we calibrate the actual positions. The target light field, L, of real camera positions is obtained through a no residual crosstalk imagination intensity distribution map for N-view hardware display system, that is two eyes receive two and only two different view images' rays at maximal simultaneously wherever (in Fig.2 (c)).

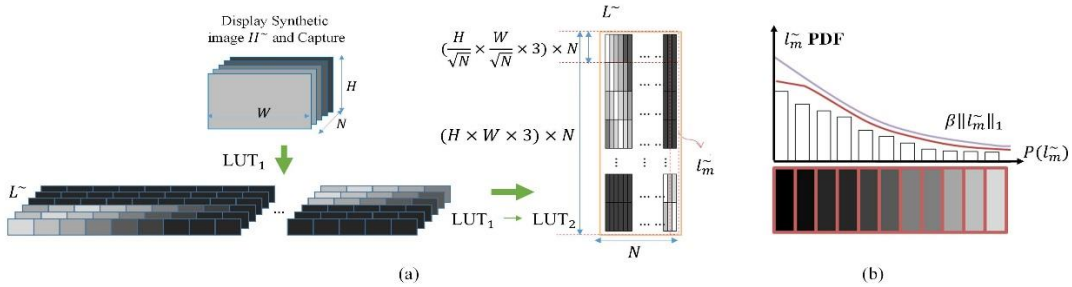

Figure 3. (a) Mapping rules from synthetic image to light field and the final representation of initial light field; (b) probability density function (PDF) of the mth point light field ($l_m^{\sim}$).

Optimization matrix: In this section, we first find the optimization matrix w, w.r.t $\min||L - wL^{\sim}||^2$. We then calculate the target synthetic image II. Since the relative physical position between the lenticular sheet and the LCD plane is unchangeable for a fixed display device, the $H \times W \times 3$ rays' optical paths from $II_{m1}{}^{\sim}$ displayed on LCD to the N capture positions are the same as the rays emitted from another point synthetic image $II_{m2}{}^{\sim}(m1 \neq m2)$. All the optical paths are calculated by the rays back-projection method [3] and a Look-Up-Table LUT_1 records the indexes from points on LCD plane to the points of viewing plane, i.e.,

$$L^{\sim} = LUT_1(II^{\sim}) \tag{1}$$

In our algorithm, the light field $L^{\sim}$ is presented as a matrix with dimension $(H \times W \times 3, N)$ (in Fig 3(a)) and the Look-Up-Table changes to LUT_2 accordingly. Each column, $l_m^{\sim}$, of $L^{\sim}$ records the light field of the mth point, and $l_{mn}^{\sim}$ $((n-1) \times H/\sqrt{N} \times W/\sqrt{N} \times 3 + 1 : n \times H/\sqrt{N} \times W/\sqrt{N} \times 3$, n from 1 to $N)$ is the photographically acquired point light field image.

The representation of L may either be exact $L = wL^{\sim}$, or approximate, $L \approx wL^{\sim}$, satisfying $min\|L - wL^{\sim}\|_2^2$. The proposed model suggests that for L the relation: $L = wL^{\sim} + v$ with a Gaussian white residual vector v with variance σ^2. w is sought that maximizes the likelihood function $P(L|w)$. l_m and $l_m^{\sim}$ represent the mth column of the light field L and $L^{\sim}$ respectively. We assume that l_m is drawn independently, readily providing: $P(L|w) = \prod_{m=1}^{N} P(l_m|w)$. Since $P(l_m|l_m^{\sim}, w) = exp\{-\|wl_m^{\sim} - l_m\|^2/2\sigma^2\}/\sqrt{2\pi\sigma^2}$, it can be computed using the following equation:

$$P(l_m|w) = \int P(l_m, l_m^{\sim}|w)\, d(l_m^{\sim})$$
$$= const \int exp\{-\|wl_m^{\sim} - l_m\|^2/2\sigma^2\} \tag{2}$$

$l_m^{\sim}$ probability was estimated from 100 light fields and a function $g(l_m^{\sim}) = \beta\|l_m^{\sim}\|_1$ was proposed to approximate the Probability Density Function (PDF) of $l_m^{\sim}$, $P(l_m^{\sim})$, the purple curve in Fig 3(b)). β is calculated by a parameter estimation method. Eq. (2) can be given by:

$$c \int exp\left\{-\frac{\|wl_m^{\sim} - l_m\|^2}{2\sigma^2}\right\}$$
$$\approx c' \int exp\{-\|wl_m^{\sim} - l_m\|^2/2\sigma^2\}\, exp\{\|l_m^{\sim}\|_1\}\, d(l_m^{\sim}) \tag{3}$$

c and c' represents two different const value. The overall problem therefore turns into:

$$arg\, max_W \sum_{m=1}^{N} max_{l_m^{\sim}} \{P(l_m, l_m^{\sim}|w)\}$$
$$= arg\, min_W \sum_{m=1}^{N} min_{l_m^{\sim}} \{\|wl_m^{\sim} - l_m\|^2 + \gamma\|l_m^{\sim}\|_1\} \tag{4}$$

An iterative method is used to solve (4), which includes two steps in each iteration: we firstly calculate the $l_m^{\sim}$ using a simple gradient descent procedure; secondly, we update the w using [8]:

$$w^{(n+1)} = w^{(n)} - \rho \sum_{m=1}^{N} l_m^{\sim T}(w^{(n)} l_m^{\sim} - l_m) \tag{5}$$

Once an optimal matrix w is calculated, the optimization synthetic image II are obtained with a reverse look up explicitly as:

$$II = LUT_2^{-1}(wL^{\sim}) \tag{6}$$

3. Result

Crosstalk measurement: The proposed method is able to improve image quality effectively. The intensity is measured in a range of viewing position. We obtained N black-and-white test images by lighting each view image in turn on the display screen [9]. Fig. 4(a) and 4(b) show the optical output of a lenticular 9-view auto stereoscopic display before and after optimization, respectively. Take the point "a" and "b" for example. In Fig. 4(a), when the second view reaches the maximum intensity of 0.97, the luminance of the first and third view are approximately 0.24 and 0.12, which causes a big crosstalk at point "a". However, our proposed method shows superior performance: both two unintended views have intensity values close to 0. This reflects that the image observed through the lenticular sheet at "b" has an equivalent effect of the second view image. Fig. 4(c) shows the result of 36-view, the red curves indicate the viewpoints and pink ones represent non-viewpoints for the 9-view hardware. Since the angle of two points for 36-view, $\alpha = \mathrm{FOV}/36 = 0.41^{\circ}$, two neighboring view images are exactly similar. Although the luminance of the first is almost 0.9 at the point "d", no evident crosstalk appears (the center panel in Fig 5.). Fig. 5 shows the display result a point "d" and "c" in Fig. 4. Table 1 compares the existing 3D synthetic image mapping systems. Experimental results validate the proposed method is simple and superior enough for high-quality 3D display productions.

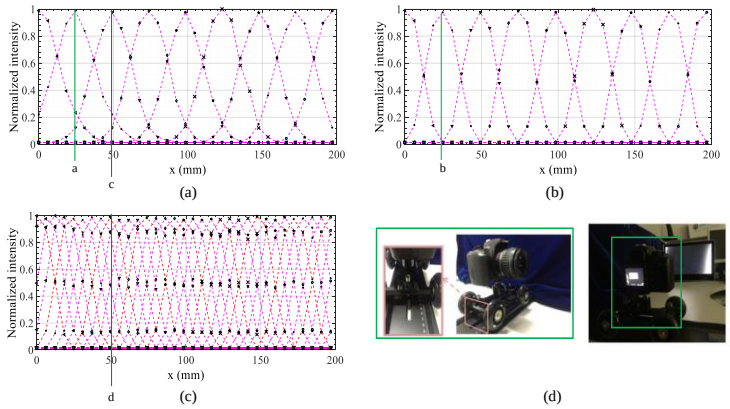

Figure 4. Visibility of different perspective views for an example lenticular multiview auto stereoscopic display when viewed from different horizontally spaced observation points, with crosstalk reduction (a) without crosstalk reduction for 9-view, (b) the proposed method for 9-view, (c) the proposed method for 36-view, (d) The experimental platform for crosstalk measurement

Figure 5. Lower left panel and enlarged right panel: without optimization, corresponding to point “d” in Fig 4.(a); Upper left panel and enlarged center panel: proposed optimized crosstalk reduction method, corresponding to point “c” in Fig 4.(c).

Table 1. Results Comparing

Method	Philips[6]	crosstalk coefficient[2]	weighted value[3]	Oversampling	subpixel multiplexing[10]	Light Field Acquisition
Computation complexity	Low	High	Low	Low	High	Low
Time for setting up a new display device	2.5h	5h	4.5h	4h	8h	3h
View number	7	9	9	27	27	36
crosstalk measurement at viewpoint	24.5%	15.4%	9.7%	23.4%	14.8%	5.2%
crosstalk at non-viewpoint	Much	Much	Much	Not too much	Less	Less
Display quality	Low	Medium	High	Medium	High	High

4. Acknowledgment

This work has been supported by project of National Natural Science Foundation of China (Grant No. 61605240) and another two projects of National High-tech R&D Program (863 Program) of China (Grant No. 2012AA011903 and Grant No. 2015AA015905) of Institute of Automation, Chinese Academy of Sciences (CASIA).

5. References

[1] Geng, Jason. "Three dimensional display technologies." Advances in optics and photonics 5.4 (2013): 456-535.
[2] Zhou, Mingcai, et al. "A unified method for crosstalk reduction in multiview displays." Journal of Display Technology 10.6 (2014): 500-507.
[3] R. Pei, Z. Geng and Z. Zhang. “A Novel Optimization Method for Lenticular 3D Display Based on Light Field Decomposion,” in Journal of Display Technology, vol. 12, no. 7, pp. 727-735, Jul. 2016. doi: 10.1109/JDT.2016.2517243.
[4] M. Levoy and P. Hanrahan. “ Light field rendering,” in Proc. 23rd Annu.Conf. Comput. Graphics and Interactive Techn., 1996, pp. 31 – 42.
[5] D. Lanman, G. Wetzstein, M. Hirsch, W. Heidrich, and R. Raskar. “ Polarization fields: Dynamic light field display using multi-layer LCDs,” Trans. Graphics ACM SIGGRAPH Asia, vol. 30, no. 6, 2011, Art. ID95.
[6] Van Berkel, Cees. "Image preparation for 3D LCD." Electronic Imaging'99. International Society for Optics and Photonics, 1999.
[7] Debevec, Paul E., and Jitendra Malik. "Recovering high dynamic range radiance maps from photographs." ACM SIGGRAPH 2008 classes. ACM, 2008.
[8] Olshausen, Bruno A., and David J. Field. "Sparse coding with an overcomplete basis set: A strategy employed by V1?." Vision research 37.23 (1997): 3311-3325.
[9] A. J. Woods. “Crosstalk in stereoscopic displays: A review,” J. Electron. Imaging, vol. 21, no. 4, Oct.–Dec. 2012, Art. ID 040902.
[10] R. Pei, Z. Geng and Z. Zhang. “ Subpixel Multiplexing Method for 3D Lenticular Display,” in Journal of Display Technology, vol. 12, no. 10, pp. 1197-1204, Oct. 2016. doi: 10.1109/JDT.2016.257647.