Three-dimensional lenticular display synthetic image rendering based on light field acquisition

Renjing Pei (SID Student Member)
Zheng Geng
Kui Ma
Mei Zhang
Rong Wang

***Abstract* —** Crosstalk is a critical defect affecting image quality in multiview lenticular 3D displays. Existing optimization methods require tedious computations and device-specific optical measurements, and results are often suboptimal. We propose a new method, on the basis of light field acquisition and optimization, for crosstalk reduction in super multiview displays. Theory and algorithms were developed, and experimental validation results showed superior performance.

Keywords *— 3D imaging, lenticular 3D display, crosstalk optimization, light field, light field capture.*

DOI # 10.1002/jsid.532

1 Introduction

Three-dimensional (3D) display technologies have been studied for decades with various ways to produce high-fidelity 3D scenes.[1] The lenticular 3D display technology, based on a sheet of cylindrical lenses (lenticulars) overlaid onto a LCD screen, is relatively simple to implement, and its quality of 3D image reconstruction appears adequate for numerous commercial 3D products and applications.[2]

The key element of lenticular 3D display is to render an accurate synthetic image, which is mapped and synthesized from mulitiview images.[3] A universal rendering rule is to determine the view number of a given point in LCD plane by Philips' method.[4] Although the implementations are simple and speedy, its algorithm often generates a 3D display with quite poor quality. 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.[5] Two different approaches are commonly used to reduce the crosstalk effects[6,7]: (1) to incorporate extra elements, for example, parallax barriers, pixel masks, or add a timing control scheme; and (2) to correct subpixel values in the synthetic images to minimize the crosstalk. Examples include crosstalk coefficient calculation[8,9] or weighted value.[3] All of these methods reduce crosstalk in the viewpoints reasonably well. However, drawbacks of these approaches include the following: (1) device-specific – they need precise measurements of optical parameters for each display device in order to maximize the crosstalk reduction; or (2) tedious and error-prone computations. Hence, the efficiency in display device calibration and 3D display performance is often suboptimal for high-quality 3D display productions.[10]

To solve this problem, we propose a light field concept to optimize lenticular 3D displays. The light field, proposed by Levoy and Hanrahan,[11] characterizes the radiance flowing through all the points in each possible direction.[12] The ultimate goal of 3D display systems is to render a "perfect" display[11,13] – 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. Theory and algorithms were developed, and experimental validation results showed superior performance (in Section 4).

2 Technical description

2.1 Overview of the proposed method

We developed algorithms for optimization based on light field framework. Experiment is conducted for multiview rendering-N-views*† ($N \in \{t | t \in Z^+, i = t/9, j = \sqrt{t}, i \in Z^+, j \in Z^+\}$) on a nine-view 3D display hardware, with the resolution of $H \times W$ and field of view of 15.3°. In this paper, the continuous light field in the real world is down-sampled to a discrete light field. We use $\tilde{II}$, II, $\tilde{L}$, L to represent the initial synthetic image, target synthetic image, emitted light field captured by camera, and target light field, respectively.

The goal is to produce the emitted light field $\tilde{L}$ and optimize it, such that the difference between emitted light field $\tilde{L}$ and the target light field L is minimal. A schematic of our technique is shown in Fig. 1. A 3D lenticular display

Received 02/07/17; accepted 02/26/17.
The authors are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, East Zhongguanchun Road, No. 95, Haidian District, Beijing 100190, China.
Renjing Pei, Kui Ma and Rong Wang are also with the School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, 101408, China; e-mail: peirenjing2013@ia.ac.cn.
*In our experiments, N = 9 or 36. When N = 36, it is called a super multiview rendering.
†M-view display hardware for N-view multiview display ($M < N$), with M viewpoint views and (N-M) non-viewpoint views.

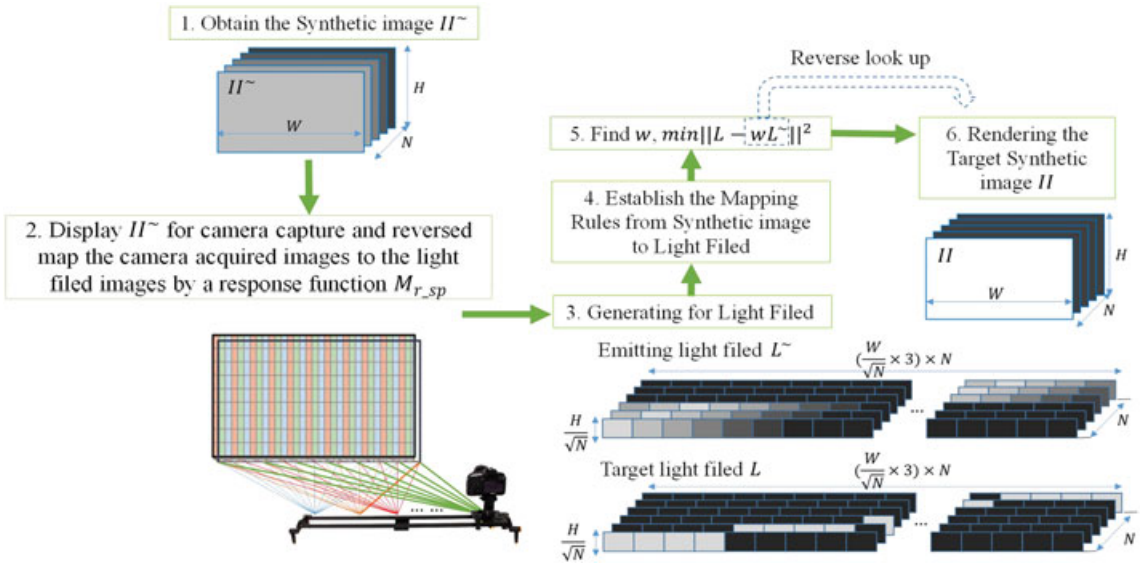


FIGURE 1 — Overview of our proposed light field decomposition method.

consists of a lenticular sheet and an LCD panel. The lenticular sheet is composed of an array of slanted cylindrical lenses. The physical structure of a cylindrical lens is shown in Fig. 2, where C represents the base plane of the cylindrical lens, p is the lenses width, and q is the distance between C and the focal plane. When a light ray (green real line) from a point on the t-axis pass through cylindrical lens, the its direction changes and becomes parallel with the line, which connects the emitting point and C (as the yellow dotted line in Fig. 2).[3] 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. 3). Correspondingly, $H/\sqrt{N} \times W/\sqrt{N} \times 3$ rays could be captured at the n-th position (assuming each pixel has three R, G, B sub-pixels). $H/\sqrt{N} \times W/\sqrt{N}$ represents the resolution of the captured image $I^{\sim}_{mn(capture)}$ $(m = 1 \sim N, n = 1 \sim N)$ photographically acquired by the camera after image preprocessing. The parameters of our display system in our experiments are listed in Table 1. In our algorithms, we render the final synthetic image on the basis of the following six-step procedures:

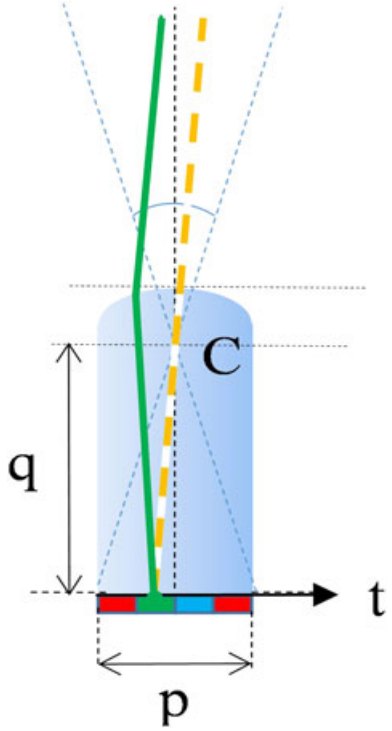


FIGURE 2 — The structure of a cylindrical lens.

(1) Obtain the initial synthetic image II: (i) Synthetic image of each point $II^{\sim}_m (m = 1 \sim N)$ is rendered when the m-th view image turned on at maximal intensity while others turned off (refer to Philips' method[4]); (ii) we obtain the initial synthetic image aggregated from $II^{\sim}_m$ as: $II^{\sim} = \cup_{m=1}^{N} II_m{}^{\sim}$ (in Fig. 4(a)). $II^{\sim}$ contains information of N points synthetic images.

(2) Capture for the light field: Display each $II^{\sim}_m$ and photographically acquire images with a moving camera recorded as $I^{\sim}_{mn(capture)}$ after capture images' preprocessed (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} similarly adopted in the literature[14] is recovered (in Section 2.4) for light field images $I^{\sim}_{mn}$.

(3) Generate the light field: (i) All the light field images in (2) are combined for an emitting light field $L^{\sim}$; (ii) The actual camera positions for each movement are exactly calibrated during the photographically acquisition in (2) (in Section 2.2), 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

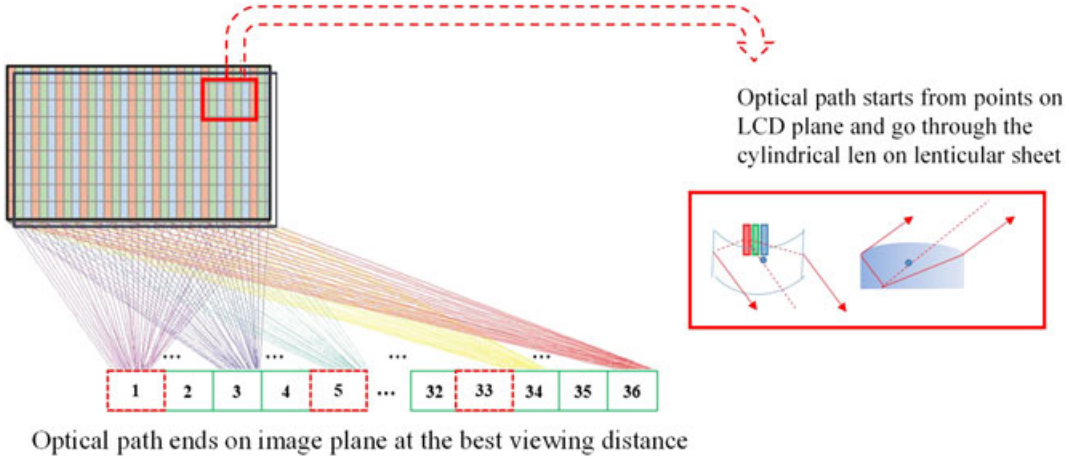


FIGURE 3 — Preset camera capture positions (N), including nine preset viewpoints marked with red squares (real line) and ($N-9$) intermediate non-viewpoints with green squares (dotted line), taking $N=36$ for example. $H\times W\times 3$ rays approximate infinite number of rays in a real light field. These rays emit from 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 n-*th* position ($n=1\sim N$).

TABLE 1 — Parameters of the 3D lenticular display in our experiment.

Parameters	Specification
Slant angel α	15.524^{o}
LCD size	21.5′
LCD resolution (H × W)	1080 × 1920
Width of cylindrical lenses p	0.5406 mm
View number	9

determine s an one-to-one mapping (recorded in a look-up-table) from points on LCD plane (synthetic image) to viewing plane by rays back-projection method in the study of Pei *et al*.[3]

(5) Compute the optimization matrix: Find the optimization matrix w: $\min\|L-w\tilde{L}\|^2$. We solve the optimization problem by a maximum likelihood method – seeking a w that maximizes the likelihood function $P(L|w)$ (in Section 2.3).

(6) Render the target synthetic image II: With an optimal matrix w, the optimization synthetic image II can be reversely searched explicitly from $w\tilde{L}$. Details are shown in following sections.

2.2 Emitting and target light field acquisitions

In this section, the acquisition of the emitting and target light field will be discussed:

2.2.1 Emitting light field

We acquire the emitting light field using a moving camera as follows: first, we determine nine viewpoints of nine-view hardware in front of the display device, at the best viewing distance. $N/9$ equal divisions are defined between two neighbor viewpoints, providing nine viewpoints and $N-9$ non-viewpoints (in Fig. 3). Then, each point synthetic image $II_m^{\sim}$ is displayed and photographically acquired 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 $N)$ is displayed sequentially, and we acquired and

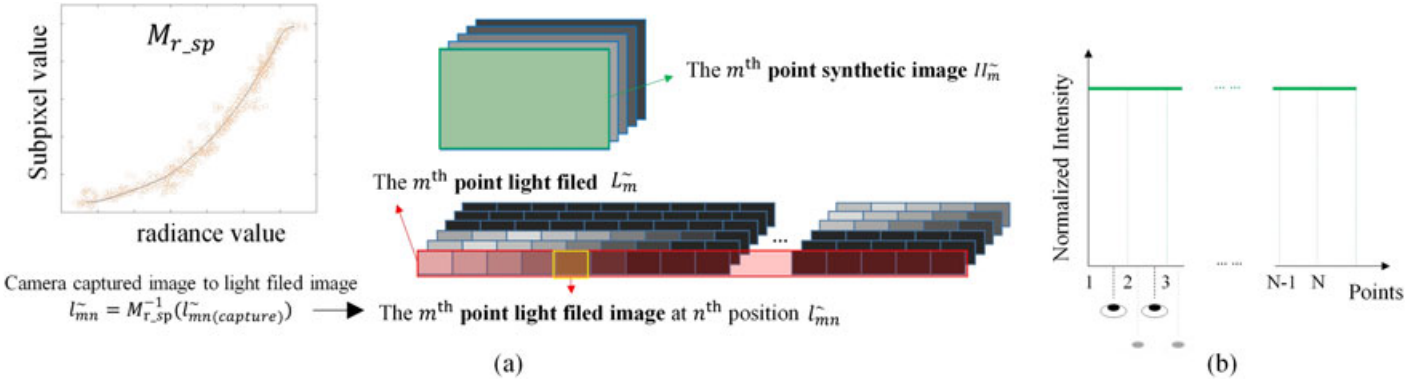


FIGURE 4 — (a) $L_m^{\sim}$ in red square denotes the m^{th} point light field, which contains N light field image $I_{mn}^{\sim}$; and $II_m^{\sim}$ in green square denotes the m^{th} point synthetic image. (b) Non-residual crosstalk imagination intensity distribution map for N-view hardware display system.

recorded each captured image as $l^{\sim}_{m1(capture)}$ (after preprocess and pre-filter. The important of image pre-filter will present in Section 3.3). The emitted light field $L^{\sim}$ is aggregated as $L^{\sim} = \cup_{m=1}^{N} L_m{}^{\sim} = \cup_{m=1}^{N} \cup_{n=1}^{N} l^{\sim}_{mn}$ and the light field image $l^{\sim}_{mn}$ (in Fig. 4(a)) can be reversely mapped by a nonlinear mapping,M_{r_sp}, from camera captured image $l^{\sim}_{mn(capture)}$ (details in Section 2.4).

We preprocess the original captured images for $l^{\sim}_{mn(capture)}$ as follows:

The interest region (the 3D display device's screen) on the undistorted captured view image is not usually a regular rectangle, so a perspective transformation is performed on the image by homograph matrix, H, which can be calculated through four corresponding pairs of points from irregular image and regular image.[15] The four points for irregular image are taken as the four corners (x_k, y_k) (k = 0, 1, 2, 3) of the 3D display device's screen, and the four points for regular image are (0, 0), $(0, W/\sqrt{N})$,$(H/\sqrt{N}, 0)$,$(H/\sqrt{N}, W/\sqrt{N})$ (in Fig. 5(b)).

2.2.2 Target light field

Because the preset positions of the camera are not precisely known during the camera movements, we calibrate the actual camera positions when we compute the target light field. A MATLAB Toolbox is used for multi-cameras' calibration. When the camera moved horizontally to capture views of the light field at each point, a calibration board was put between 3D display device and camera also captured to facilitate the camera position calibration (in Fig. 5(a)).

The target light field, L, of real camera positions is obtained through a non-residual crosstalk imagination intensity distribution map for N-view hardware display system. Two eyes receive two and only two different view images' rays at maximal simultaneously wherever (in Fig. 4(b)).

2.3 Optimization matrix *w*

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.

Because 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 $II^{\sim}_{m1}$ displayed on LCD to the N capture positions are the same as the rays emitted from another point synthetic image $II^{\sim}_{m2}(m1 \neq m2)$. All the optical paths are calculated by the rays back-projection method (in Fig. 3),[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. 6(a)) and the Look-Up-Table changes to LUT_2 accordingly. Each column, $l^{\sim}_m$, of $L^{\sim}$ records the light field of the m-th point, and $l^{\sim}_{mn}$ $((n-1).\times H/\sqrt{N}\times W/\sqrt{N}\times 3 + 1{:}n\times H/\sqrt{N}\times W/\sqrt{N}\times 3$, 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 $l^{\sim}_m$ represent the m-th column of the light field L and $L^{\sim}$, respectively. We assume that l_m is drawn independently, readily providing: $P(L|w) = \prod_{m=1}^{N} P(l_m|w)$. Because $P(l_m|l^{\sim}_m, w) = \exp\left\{-\left\|wl^{\sim}_m - l_m\right\|^2/2\sigma^2\right\}/\sqrt{2\pi\sigma^2}$, it can be computed using the following equation:

$$\begin{aligned} P(l_m|w) &= \int P(l_m, l^{\sim}_m|w)\, d(l^{\sim}_m) \\ &= const \int \exp\{-wl^{\sim}_m - l_m{}^2/2\sigma^2\} \end{aligned} \tag{2}$$

$l^{\sim}_m$ probability was estimated from 100 light fields, and a function $g(l^{\sim}_m) = \beta l^{\sim}_m 1$ was proposed to approximate the probability density function of $l^{\sim}_m$, $P(l^{\sim}_m)$, the purple curve in Fig. 6(b)). β is calculated by a parameter estimation method. Eq. (2) can be given by

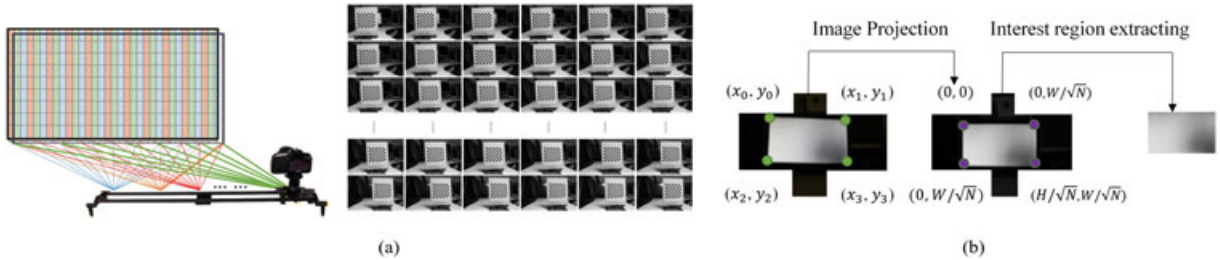


FIGURE 5 — (a) A calibration board is captured for camera positions calibration during the light field acquisition process. (b) Interest region (the 3D display device's screen) extraction.

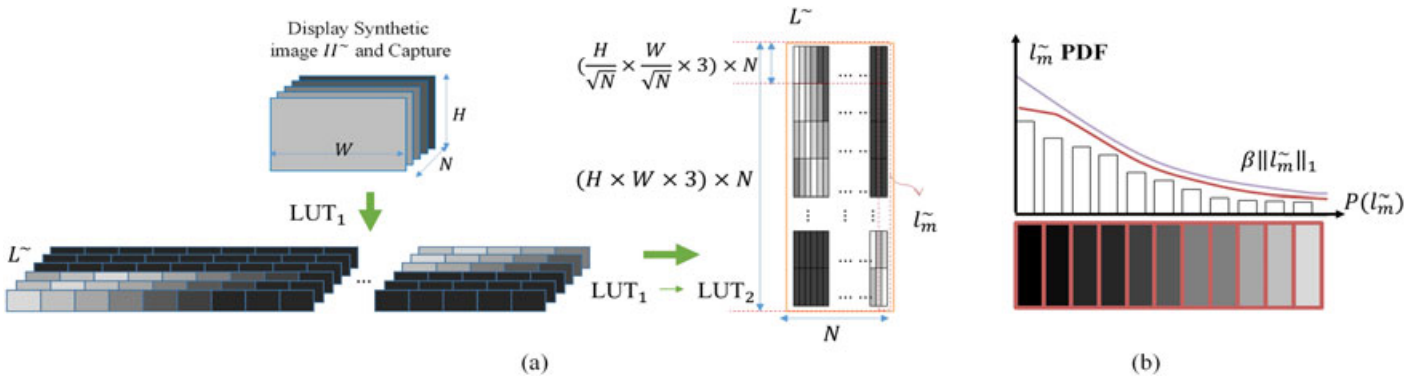


FIGURE 6 — (a) Mapping rules from synthetic image to light field and the final representation of initial light field; (b) probability density function (PDF) of the *m*-th point light field ($l_m^{\sim}$).

$$c\int \exp\{-wl_m^{\sim} - l_m{}^2/2\sigma^2\} \approx c'\int \exp\{-wl_m^{\sim} - l_m{}^2/2\sigma^2\} \exp\{l_m^{\sim}1\} d(l_m^{\sim}) \tag{3}$$

c and c' represent two different const values. The overall problem therefore turns into

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

An iterative method is used to solve Eq. (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[16]

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

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

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

2.4 M_{r_sp} mapping function fitting

This section presents our algorithm for recovering the M_{r_sp} response function from radiance value (light field image $l_{mn}^{\sim}$) to subpixel′ value on acquired image $l_{mn(capture)}^{\sim}$; that is, $l_{mn(capture)}^{\sim} = M_{rsp}(l_{mn}^{\sim})$.

The fitting algorithm is similar to the algorithm in the study of Debevec and Malik.[14] We take a number of digitized photographs from the same scene with same exposure (the lighting changes can be safely ignored). After the camera scanning and digitization processes, we obtain a digital number z_i for each subpixel on acquired image. We may write the nonlinear function, M_{r_sp}, of the original radiance value, l_i, at the subpixel as

$$z_i = M_{r\_sp}(l_i) \tag{7}$$

We assume that M_{r_sp} is smooth and i ranges over subpixels of the digitized photographs. In this set of equations, z_i and l_i are known and the function M_{r_sp} is unknown. We wish to recover the function M_{r_sp} that best satisfy the set of equations of Eq. (7) in a least-squared error sense, so our goal is to minimize the following quadratic objective function:

$$E = \sum_{i=1}^{N} \left[z_i - M_{r\_sp}(l_i)\right]^2 + \alpha \sum_{i=1}^{N} M_{r_{sp}}{}''(l_i), \tag{8}$$

where N is the number of subpixel. The scalar α weights the smoothness term relative to the data fitting term. We use a three-parameter function $t(x) = \alpha + \beta x^{\gamma}$ to characterize the response function M_{r_sp}.

Because the subpixel values in the middle of their range are more reliable and the function M_{r_sp} becomes monotonic in the vicinity of its saturation value, we add a weighted function $\omega(z)$, which decays to zero at both ends of the subpixel value range.

$$w(z) = \begin{cases} z - z_{min} & z \le (z_{min} + z_{max})/2 \\ z_{max} - z & z > (z_{min} + z_{max})/2 \end{cases}$$

where $z_{min} = 0$ and $z_{max} = 255$ in our experiments. We modify Eq. (8) to a weighted energy function as

$$\begin{aligned} E = &\sum_{i=1}^{N} \gamma_1 \times w(z_i) \times \left[z_i - M_{r\_sp}(l_i)\right]^2 \\ &+ \alpha \sum_{i=1}^{N} \gamma_2 \times w(z_i) \times M_{r_{sp}}{}''(l_i) \end{aligned} \tag{9}$$

γ_1 and γ_2 are the scalar terms. The final fitting result is shown in Fig. 7.

3 Results and discussion

3.1 Crosstalk measurement

The method we proposed is able to improve image quality effectively. A camera was moved horizontally to measure the intensity in a range of viewing position at a predetermined viewing distance. We obtained N black-and-white test images by lighting each view image in turn on the display screen. For these tests, only subpixels corresponding to a certain view image were turned on at maximal intensity and all other subpixels were turned off. Then, we measured each captured image's intensity, respectively, along the horizontal direction.[17]

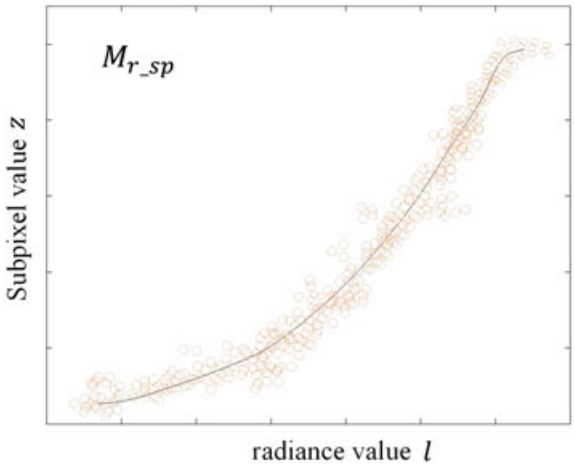


FIGURE 7 — Recovered response curves for response function M_{r_sp}.

Figure 8(a) and (b) show the optical output of a lenticular nine-view auto stereoscopic display before and after optimization, respectively. Take the point "a" and "b" for example. In Fig. 8(a), when the second view reaches the maximum intensity of 0.97, the luminance of the first and third view is approximately 0.24 and 0.12, which causes a significant crosstalk at point "a". However, our proposed method shows superior performance: Both two unintended views have intensity values close to 0. This demonstrates that the image observed through the lenticular sheet at "b" has an equivalent effect of the second view image. Figure 8(c) shows the result of 36-view, the red curves indicate the viewpoints, and pink ones represent non-viewpoints for the nine-view hardware. Because 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. 9).

3.2 View number for mapping

The commonly used metric mean squared error was used to evaluate the result, $\mathrm{MSE} = L - \mathrm{w}L\tilde{}_2^2/(N\times H\times W\times 3)$, which was converted to the peak signal-to-noise ratio (PSNR) = $10\log_{10}(255^2/\mathrm{MSE})$. Larger the PSNR implies L is closer to $\mathrm{w}L\tilde{}$. It is obvious that the more views in rendering the closer the light field generated is to that produced by real-world physical objects. The emitted light

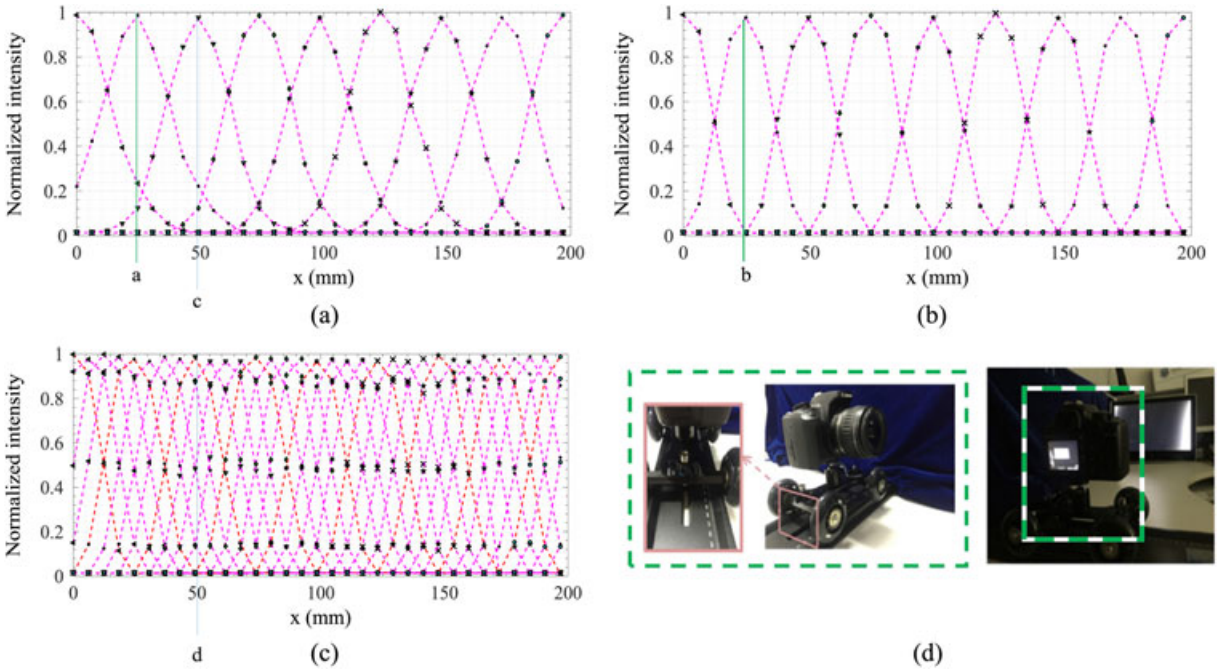


FIGURE 8 — Visibility of different perspective views for an example lenticular multiview auto stereoscopic display when viewed from different horizontally spaced observation points, with crosstalk reduction (a) without crosstalk reduction for nine-view, (b) the proposed method for nine-view, (c) the proposed method for 36-view, (d) the experimental platform for crosstalk measurement.

FIGURE 9 — Lower left panel and enlarged right panel: without optimization, corresponding to point "d" in Fig. 8(a); Upper left panel and enlarged center panel: proposed optimized crosstalk reduction method, corresponding to point "c" in Fig. 8(c).

TABLE 2 — Mapping view number comparing.

Number of mapping views (N)	9	36
PSNR	30.75	45.62

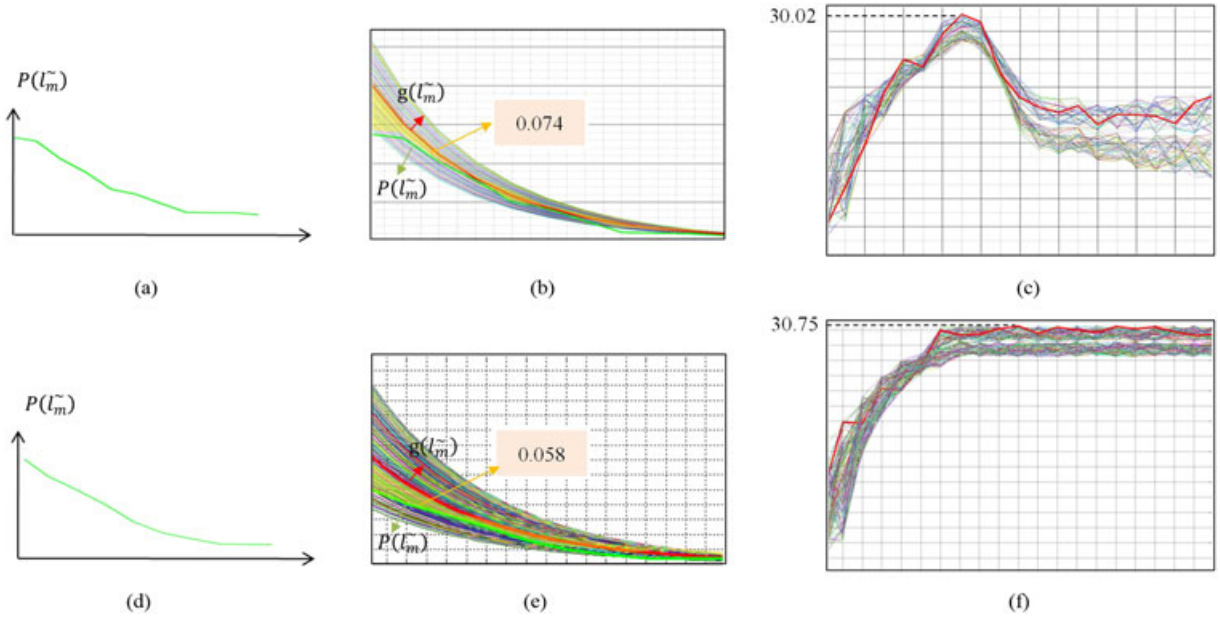


FIGURE 10 — (a), (b), and (c) (/(d), (e) and (f)) are the intermediate results without (/with) Gaussian filter. The green curves in (a) and (b) (/(d) and (e)) are the probability density function of $\tilde{l_m}$. The other color (except for green) curves in (b), (c) (/(e), (f)) illustrate approximate functions of $g(\tilde{l_m})$ with different parameter β, and the red curve represents the function $g(\tilde{l_m})$ approximating the probability density function of $\tilde{l_m}$ in our experiments.

TABLE 3 — Comparison of results.

Method	Philips[4]	Crosstalk coefficient[8]	Weighted value[3]	Oversampling	Subpixel multiplexing[18]	Light field acquisition
Computation complexity	Low	High	Low	Low	High	Low
Time for setting up a new display device	2.5 h	5 h	4.5 h	4 h	8 h	3 h
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

field thus has higher similarity to the target light field, as shown in Table 2. Figure 9 shows the display result a point "d" and "c" in Fig. 8.

3.3 The importance of image $l_{mn}^{\sim}$ pre-filter

Difference in image qualities of photographical image acquisition results in diverse optimization light field results, and a pre-filter, like the Gaussian filter, is essential for the optimization. Figure 10(b) and (e) shows the estimation result for function $g(l_m^{\sim})$ in optimization process. Figure 10(c) and (f) presents the PSNR values with different iterations without and with image preprocessing, respectively. Because of the image noise influence, PSNR fails to reach final convergence. Taking $N = 9$ for example, when $n = 105$, $PSNR = 30.02$. But as the iteration number n continually increases, PSNR declines again as the Fig. 10(c) shown. After applying Gaussian filter on acquired images $l_{mn}^{\sim}$, we were able to obtain the maximum value 30.75 of the PSNR and a converged result. The area of the yellow polygon S in Fig. 10(b) and (e) represents the differences between $g(l_m^{\sim})$ and $P(l_m^{\sim})$. The smaller the difference is, the closer the optimized light field converges to the target light field. S is approximately 0.058 and 0.074 with and without the Gaussian filter, respectively.

4 Impact

Table 3 compares the existing 3D synthetic image mapping systems. Philips Research Laboratories[4] adopted seven-view architecture to populate individual views that reach the observer's eyes. Zhou[8] developed a method to correct subpixel values in the synthetic images, for example, crosstalk coefficient calculation. Pei[3] considered all arriving rays to derive each subpixel's weighted value to alleviate crosstalk and smooth fusion of different views and the subpixel multiplexing method for 3D lenticular display.[18] We modified view number parameter in Philips to obtain oversampled synthetic image. We measure the crosstalk as the Section 3.1 mentioned – with N black-and-white test images by lighting each view image in turn on the display screen. Crosstalk can be finally mathematically define as

$$\text{Crosstalk} = leakage/(signal + leakage) \times 100\%,$$

where "leakage" is the luminance of light that leaks from the unintended channel to the intended channel and "signal" is the luminance of the intended channel. Experimental results validate that the proposed method is simple and superior enough for high-quality 3D display productions.

Acknowledgments

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).

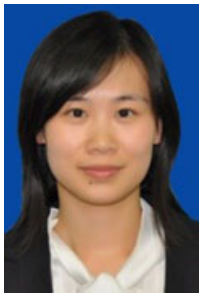
References

1 T Hsu *et al.*,"High resolution autostereoscpic 3-D display with proximity projector array," in Proc. SID, 2008, pp. 760–763.
2 T. Okoshi, 3 Dimensional Imaging Techniques. Academic, New York, (1976) ch. 2, pp. 8–42.
3 R. Pei *et al.*, "A novel optimization method for lenticular 3D display based on light field decomposion," *J. Disp. Technol.*, **12**, No. 7, 727–735 (2016) DOI:10.1109/JDT.2016.2517243.
4 V. Berkel, "Cees. "Image preparation for 3D LCD." Electronic Imaging'99". International Society for Optics and Photonics, (1999).
5 J. Geng, "Three-dimensional display technologies," *Adv. Opt. Photon.*, **5**, No. 4, 456–535 (2013).
6 C. Lee *et al.*, "Auto-stereoscopic 3D displays with reduced crosstalk," *Opt. Express*, **19**, No. 24, 24762–24774 (2011).
7 L. Xing *et al.*, "Assessment of stereoscopic crosstalk perception," *IEEE Trans. Multimedia*, **14**, No. 2, 326–337 (2012).
8 M. Zhou *et al.*, "A unified method for crosstalk reduction in multiview displays," *J. Disp. Technol.*, **10**, No. 6, 500–507 (2014).
9 X. Wang and C. Hou, "Improved crosstalk reduction on multiview 3D display by using BILS algorithm," *J. Appl. Math.*, **2014**, Article ID 428602 (2014).
10 X. Li *et al.*, "Image processing to eliminate crosstalk between neighboring view images in three-dimensional lenticular display," *J. Disp.technol.*, **7**, No. 8 (2011).
11 M. Levoy and P. Hanrahan, "Light field rendering," in Proc. 23rd Annu. Conf. Comput. Graphics and Interactive Techn., 31–42 (1996).
12 D. Lanman *et al.*, "Polarization fields: dynamic light field display using multi-layer LCDs," *Trans. Graphics ACM SIGGRAPH Asia*, **30**, No. 6 (2011) Art. ID95.
13 A. Jones *et al.*, "Rendering for an interactive 360° light field display," in SIGGRAPH 2007 Papers (2007), paper 40.
14 P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs." ACM SIGGRAPH 2008 classes. ACM, (2008).
15 P. H. S. Torr and A. Zisserman, "MLESAC: a new robust estimator with application to estimating image geometry," *Comput. Vis. Image Underst.*, **78**, No. 1, 138–156 (2000).
16 B. A. Olshausen and D. J. Field, "Sparse coding with an overcomplete basis set: a strategy employed by V1," *Vision Res.*, **37**, No. 23, 3311–3325 (1997).
17 A. J. Woods, "Crosstalk in stereoscopic displays: a review," *J. Electron. Imaging*, **21**, No. 4, 040902 (2012) Art. ID.
18 R. Pei *et al.*, "Subpixel multiplexing method for 3D lenticular display," *J. Disp. Technol.*, **12**, No. 10, 1197–1204 (2016) DOI:10.1109/JDT.2016.2576471.

Renjing Pei received the BS degree in computer science from the WuHan University, China, in 2013. She is currently working toward the PhD degree in control theory and control engineering at the institute of automation, Chinese academy of sciences, Beijing, China. Her research interests are 3D display and 3D reconstruction. Fulfilling periodical paper research and publishing and related program tasks with good quality and quantity in time. Two SCI index papers and four international conference papers have been published, one of which was nominated as the excellent papers and was recommended to journal (6%).

Zheng Geng received his doctoral degree in Electrical Engineering from George Washington University in 1990. Since then, he has led a variety of research, development, and commercialization efforts on 3D imaging technologies. He founded a high tech company that specialized in developing 3D imaging technologies and products. He has served as a principal investigator and major contributor for over $35 million in research and product efforts. He has published 142 academic papers and one book, and is an inventor of 33 issued patents. He has received prestigious national honors, including the Tibbetts Award from the U.S. Small Business Administration and the "Scientist Helping America" award from the Defense Advanced Research Projects Agency, and was ranked 257 in INC. magazine's "INC. 500 List." Dr. Geng currently serves as the vice-president for the IEEE Intelligent Transportation Systems Society (ITSS). He is also leading Intelligent Transportation System standard efforts by serving as the chairman of the standard committee for IEEE ITSS. His recent publication on 3D imaging technology ("Structured light 3D surface imaging: a tutorial," www.opticsinfobase.org/aop/abstract.cfm?uri=aop-3-2-128) was among the top downloaded articles in the OSA Advances in Optics and Photonics (AOP) journal.

Kui Ma received the BS degree in automation from the Harbin Engineering University, China, in 2015. He is currently working toward the MS degree in control theory and control engineering at the institute of automation, Chinese academy of sciences, Beijing, China. His research interests are 3D display.



Mei Zhang is a research assistant in the Institute of Automation of Chinese Academy of Sciences, Beijing, China. She received her PhD in optical engineering from the Nankai University, China, in 2010. Her work focuses on the research of physics principles and optical design method. In particular, she is interested in the technology development for the 3D light field endoscopy and 3D light field system. She has published near 20 papers in multiple journals.

Rong Wang received the BS degree in automation from the Beijing University of Aeronautics and Astronautics, China, in 2013. She is currently working toward the PhD degree in computer application at the institute of automation, Chinese academy of sciences, Beijing, China. Her research interest is computer visual.