High-rank Coded Aperture Projection for Extended Depth of Field

Chenguang Ma[1], Jinli Suo[1], Qionghai Dai[1], Ramesh Raskar[2], Gordon Wetzstein[2]

[1]Department of Automation, Tsinghua University, [2]MIT Media Lab

Abstract

Projectors require large apertures to maximize light throughput. Unfortunately, this leads to shallow depths of field (DOF), hence blurry images, when projecting on non-planar surfaces, such as cultural heritage sites, curved screens, or when sharing visual information in everyday environments. We introduce high-rank coded aperture projectors—a new computational display technology that combines optical designs with computational processing to overcome depth of field limitations of conventional devices. In particular, we employ high-speed spatial light modulators (SLMs) on the image plane and in the aperture of modified projectors. The patterns displayed on these SLMs are computed with a new mathematical framework that uses high-rank light field factorizations and directly exploits the limited temporal resolution and contrast sensitivity of the human visual system. With an experimental prototype projector, we demonstrate significantly increased DOF as compared to conventional technology.

1. Introduction

Projectors have evolved into a ubiquitous technology. Today, these devices are essential for many different applications and they come in a wide range of form factors, including smartphones and other hand-held devices, integrated in offices, conference rooms, large-scale installations, and movie theaters. A major trend for projectors is that they are being used in everyday environments, without special canvases, to share information and collaborate using a common workspace. Two of the biggest technical demands on projection displays being used for such applications are image sharpness and brightness on whatever surface they are projected on. High brightness is achieved through large apertures; unfortunately, this results in shallow depths of field, therefore, reduced sharpness. In this paper, we introduce a new approach to extended depth of field (DOF) projection that facilitates sharp and light efficient projections on a variety of different surfaces.

For this application, we propose a computational display approach. Our method combines high-speed coded projec-
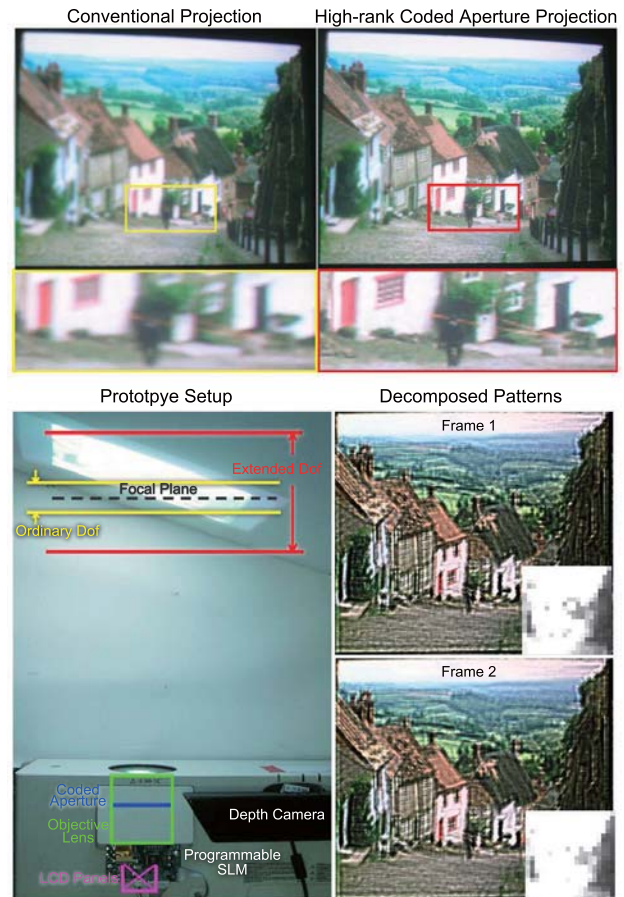

Figure 1. High-rank coded aperture projection system. We show that high-speed programmable apertures in combination with a new light field factorization algorithm can extend the depth of field of a projection display. The upper row shows projections on a non-planar surface with a conventional projector (left) and our high-rank display (right). Image quality is significantly enhanced. The lower row shows our prototype device with optical elements indicated (left) and also two of the decomposed patterns for the coded aperture and the image SLM inside the projector (right).

tor apertures with novel computational routines that solve a light field factorization problem to generate the patterns

necessary for high-quality extended DOF display. With our mathematical framework, we build on recent developments in compressive light field displays [9, 22]. While these approaches synthesize high-rank light fields for the purpose of glasses-free 3D image display, we synthesize high-speed image and aperture patterns for spatial light modulators (SLMs) inside a projector with the goal of enhancing the depth of field of such a device.

In particular, we make the following contributions:

- We introduce high-rank coded aperture projectors as a new computational display technology that employs high-speed SLMs and a novel, projected light field factorization algorithm.

- We demonstrate practical application of our computational display architecture to extended depth of field projection.

- We evaluate the proposed system with a hardware prototype and with quantitative simulations.

2. Related Work

Extended Depth of Field Photography   has been a very active area of research for more than a decade. In most of the proposed approaches, the aperture of a camera is optically modulated with phase [7, 12, 5] or intensity [11, 20, 24] codes. In a computational reconstruction step these codes are inverted, usually using deconvolution, and the depth of field of a captured image extended. Inspired by ideas of extending DOF in photographs, we explore coded apertures for extended depth of field projection with a new computational display architecture.

Extended Depth of Field Projection   has also been investigated. Existing methods can be categorized into approaches that modify conventional projector optics and those that do not. The latter category includes the work by Brown et al. [4], Oyamada and Saito [17], and Zhang and Nayar [23]; in all of these systems, the depth-dependent blur of the projection on a known surface is pre-corrected in software. It is well known, however, that circular or Gaussian blur kernels, as exhibited by conventional projectors, optically and irreversibly cancel out high spatial frequencies. Therefore, the performance of purely software-based methods is fundamentally limited by the available optics. Bimber and Emmerling [2] proposed to overlay the images of multiple projectors, each focused at a different depth; a composite image is then computed that minimizes the overall blur of the resulting projection. More recently, Grosse et al. [8] introduced single-device coded aperture projection techniques that jointly optimize the aperture code and the projected image for the application of extended DOF display. Our method is closely related to their approach;

however, we employ high-speed SLMs in combination with a novel, high-rank projected light field decomposition that allow us to significantly extend the achieved depth of field.

High-rank Light Field Displays   have been explored for the application of glasses-free 3D displays. Lanman et al. [9] showed that the light field emitted by a duallayer LCD display can be parameterized as a matrix. By temporally displaying patterns at a speed that is higher than the flicker fusion threshold of the human visual system, they are able to synthesize high-rank light fields. This idea was extended to a variety of optical setups, including multilayer displays and directional backlighting by Wetzstein et al. [22]. These methods employ nonnegative matrix and tensor factorizations that decompose a target light field into a set of patterns shown on the respective displays. In this paper, we show that a coded aperture projector also emits a light field; this is, however, down-projected from 4D spatioangular light field space onto 2D screen space. We formulate the image formation and coded aperture synthesis as a light field factorization problem and demonstrate that it performs better than previous models.

3. High-rank Coded Aperture Projection

In this section, the image formation of projection is introduced and interpreted as an angular integration of the light field emitted by the projector. With time-multiplexed high-rank approximation, nonnegative matrix factorization minimizes the weighted Euclidean distance between a target projection image and the above integration. Furthermore, perceptual measurement is also incorporated into the optimization which leads to a better visual quality of the projected imagery.

3.1. Image Formation

Ordinary projectors have large aperture stops to optimize light transmission, and therefore shallow depths-of-field. When displaying imagery onto out-of-focus or non-planar screens, the projected target image will be blurred due to optical defocus. This optical image formation is traditionally described as a convolution of the original image with a filter kernel that corresponds to the aperture of the projector, and the scale of the kernel is directly proportional to the distance between the screen and the focal plane.

Our proposed projectors can be cast as a dual-layer light field emitting system, with two light-attenuating layers illuminated by a bright uniform back light: one on the image plane inside the projector and the other on the aperture plane in the objective lens. For our analysis, we adopt a variant of the two-plane parameterizations [13] of the 4D light field. As shown in Figure 2, an emitted ray is parameterized by its intersection coordinates with both mask layers. Unlike
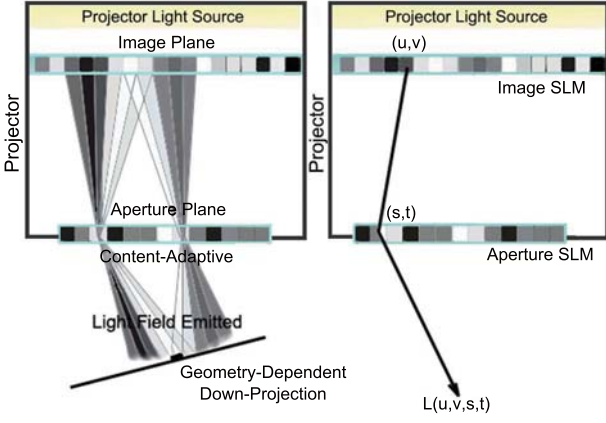
Figure 2. Principle of projected light field. We consider dual-layer patterns (one on the image plane and the other on the aperture plane) as general spatial light modulators that act in concert to recreate a target light field and the projection on the screen is viewed as an geometry-dependent down-projection.

conventional two-plane parameterizations, the rays here are refracted by the projection lenses and thus deflected at the front mask on the aperture plane. We denote that the ray (u, v, s, t) intersects the image SLM at the point (u, v) and the aperture SLM at the point (s, t). Thus the emitted light field can be given as

$$\boldsymbol{L}(u, v, s, t) = \boldsymbol{f}(u, v) \otimes \boldsymbol{g}(s, t). \tag{1}$$

For simplification, we assume the screen to be diffuse surface with constant albedo. Then with the definition of light field, the formation of an in-focus projection is an integration over the angular light field dimension:

$$\boldsymbol{i}_{d=d_f}(x, y) = \iint_{\Omega_{st}} \boldsymbol{L}(u, v, s, t) ds dt \tag{2}$$

where $\boldsymbol{i}_d(x, y)$ is the displayed image at distance d with coordinate (x, y), d_f is the distance of focal plane. For out-of-focus projections, however, the integration within the spatial-angular light field will be not just over the angular dimension but an integration over a screen geometry-dependent manifold:

$$\boldsymbol{i}_d(x, y) = \iint_{\Omega_{st}} \boldsymbol{L}(\boldsymbol{\phi}(x, y, d), s, t) ds dt \tag{3}$$

where $\boldsymbol{\phi}(x, y, d)$ returns the relevant image plane coordinates of projected pixel (x, y) at distance d. Then using a delta function, we can formulate Equation (3) as

$$\boldsymbol{i}_d(x, y) = \iint_{\Omega_{st}} \iint_{\Omega_{uv}} \boldsymbol{\delta}((u, v) - \boldsymbol{\phi}(x, y, d)) \boldsymbol{L}(u, v, s, t) du dv ds dt \tag{4}$$

Now, every pixel of the image is an accumulation of light rays from different directions and we can view the projected imagery as an integration of the emitted light field. In other words, the projection screen works as a geometry-dependent 2D down-projection of the 4D light field, which can be further expressed as matrix multiplication.

3.2. Projected Light Field Formulation

We will show in this subsection that the projected light field can be formulated as a matrix multiplication for optimization. The pair of 2D masks is discretely defined as $\boldsymbol{f}[u, v]$ and $\boldsymbol{g}[s, t]$, corresponding to the rear and front layers. As shown in Figure 2, the emitted 4D discretized light field is defined as $\boldsymbol{L}[u, v, s, t]$, corresponding to the continuous coordinates (u, v, s, t), and is given by the outer product of the masks as

$$\boldsymbol{L}[u, v, s, t] = \boldsymbol{f}[u, v] \otimes \boldsymbol{g}[s, t]$$
$$\boldsymbol{L} = \boldsymbol{g} \boldsymbol{f}^T \tag{5}$$

The down-projection of the emitted light field on the screen depends on the geometry of the projection screen. As shown in Figure 3, light rays from each projector pixel on an out-of-focus screen correspond to different pixels on the rear SLM, and is directly related to the distance between the screen and the projector's focal plane. We use a Microsoft Kinect to capture the screen geometry. In this way, we measure the spatially varying depth information for the geometry-dependent down-projection. This is a one-time calibration step and can be done by any other depth estimation, such as projector-camera system based structured illumination.

As shown in Figure 3, the geometry-dependent down-projection at a certain depth d (planar screen) can be seen as a summation of sampling elements from different columns of the light field (marked in red). Written as matrix multiplication, this down-projection of Equation (4) can be formulated as

$$\boldsymbol{i}_d(x, y) = \sum_{i=1}^{n} \boldsymbol{p}_{i,d} \boldsymbol{L} \boldsymbol{Q}_i = \sum_{i=1}^{n} \boldsymbol{p}_{i,d} \boldsymbol{g} \boldsymbol{f}^T \boldsymbol{Q}_i \tag{6}$$

where n is the number of pixels involved in the circle of confusion on image plane when the projection is defocused at depth $\boldsymbol{d}$, $\boldsymbol{Q}_i$ is column-shifted matrix for the ith involved image plane pixel and $\boldsymbol{p}_{i,d}$ is row-sampling matrix representing the pixels on the aperture plane through which the light ray of pixel i travels. For a projection of resolution $h \times w$ and a coded aperture of resolution $m \times n$, the dimensions of the matrices are: $\boldsymbol{f}$ $hw \times 1$, $\boldsymbol{g}$ $mn \times 1$, $\boldsymbol{p}_{i,d}$ $1 \times mn$ and $\boldsymbol{Q}_i$ $hw \times hw$.

Furthermore, to deal with complex geometry where $\boldsymbol{d}$ is of varying values (non-planar screen), we first discretize the depth map of the screen and divide it into the t depth
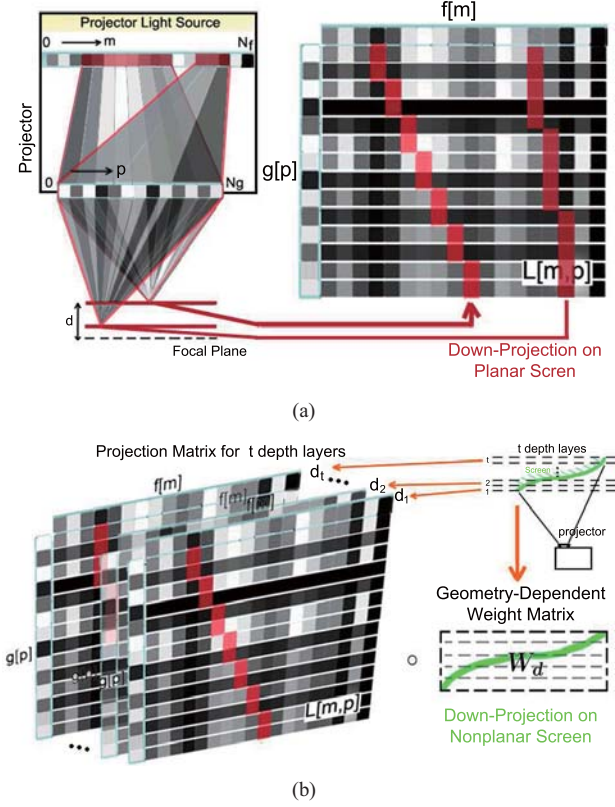
Figure 3. Derivation of geometry-dependent down-projection on 2D slice of the 4D light field. (a) Down-projection on planar screen is derived as a summation of sampling elements (indicated in red). (b) In non-planar case, the depth map of the screen is divided into t layers and represented as the nonzero elements in the weight matrix $\boldsymbol{W_d}$ (in green), then by Hadamard product with the projection matrix consists of planar down-projections at different distances, down-projection on non-planar screen is formulated.

layers, as shown in Figure 3. After that, we build the projection matrix for each depth layer by using the similar column-shifted matrices $\boldsymbol{Q_i}$ and the row-sampling matrices $\boldsymbol{p_{i,d}}(d = 1, 2...t)$ which merge up to be $\boldsymbol{P_i}$ with a dimension of $t \times mn$. Also we expand n to be the maximum number of pixels in the circle of confusion on the image plane when the projection is defocused at the largest depth of $\boldsymbol{d_t}$. Then we integrate the depth map of the screen by adding a geometry-dependent weight matrix $\boldsymbol{W_d}$ with a dimension of $t \times hw$, which has nonzero elements w_{tp} corresponding to a screen pixel p at depth layer t as shown in Figure 3. Using Hadamard product between the layer stack and the geometry-dependent weight matrix, down-projection of the light field on non-planar screen can be formulated as

$$\boldsymbol{i}_d(x, y) = \boldsymbol{W_d} \circ \sum_{i=1}^{n} \boldsymbol{P_i} \boldsymbol{g} \boldsymbol{f}^T \boldsymbol{Q_i}. \qquad (7)$$

where $\circ$ denotes the Hadamard product for element-wise multiplication of matrices.

3.3. Optimizing Image and Aperture Patterns

Our approach of defocus compensation can be simplified as matrix approximation based on projected light field decomposition. We implement a projected variant of nonnegative matrix factorization to solve for the image and aperture patterns. Displayed at high-speeds, these patterns will be averaged in the human visual system, creating a high-rank approximation of the projected light field.

Following the concept of temporal multiplexing for parallax barriers in [9], we consider all possible mask pairs for a sequence of dual-layer patterns. With higher speed SLMs, more degrees of freedom can be gained from the temporal domain. Specifically, the light field emitted by the time-multiplexed dual-layer SLMs can be decomposed as the matrix product

$$\boldsymbol{L}[u, v, s, t] = \sum_{k=1}^{T} \boldsymbol{f}[u, v] \otimes \boldsymbol{g}[s, t] = \boldsymbol{G}\boldsymbol{F}^T \qquad (8)$$

where $\boldsymbol{F}$ and $\boldsymbol{G}$ are of size $N_f \times T$ and $N_g \times T$ respectively. N_f and N_g are the size of the image and the aperture. Column k of $\mathbf{F}$ and $\mathbf{G}$ are the masks displayed on the rear and front SLMs layers at kth frame, respectively.

By our derivation of projected light field formulation, together with the intrinsic constraint of non-negativity, the solution of searching the best content-adaptive layer patterns can be given by seeking a non-negative matrix factorization $\boldsymbol{G}\boldsymbol{F}^T$ that minimizes the weighted Euclidean distance between target sharp image and angular integrated projection, i.e.,

$$\underset{\boldsymbol{F}, \boldsymbol{G}}{\operatorname{argmin}} \|\boldsymbol{I} - \sum_{i=1}^{n} \boldsymbol{P_i} \boldsymbol{G} \boldsymbol{F}^T \boldsymbol{Q_i}\|_{\boldsymbol{W_d}}^2, s.t. \boldsymbol{F}, \boldsymbol{G} \geq 0. \qquad (9)$$

Here $\boldsymbol{F}$, $\boldsymbol{G}$ are the sequential SLM patterns at the image and aperture plane, $\boldsymbol{I}$ is the target image stack, and $\boldsymbol{P_i}$, $\boldsymbol{Q_i}$, $\boldsymbol{W_d}$ is the same as defined above with the dimensions of $t \times N_g$, $N_f \times N_f$ and $t \times N_f$ respectively.

Non-negative matrix factorization can be used to solve Equation (9). Prior numerical methods include the multiplicative update rule [10] and the weighted update [1][3]. We modify the weighted update algorithm and use the multiplicative update rules as follows.

$$
\begin{aligned}
\boldsymbol{G} &\leftarrow \frac{\boldsymbol{G} \circ \sum_i \boldsymbol{P_i}^T (\boldsymbol{W_d} \circ \boldsymbol{I}) \boldsymbol{Q_i}^T \boldsymbol{F}}{\sum_i \boldsymbol{P_i}^T (\boldsymbol{W_d} \circ \sum_j \boldsymbol{P_j} \boldsymbol{G} \boldsymbol{F}^T \boldsymbol{Q_j}) \boldsymbol{Q_i}^T \boldsymbol{F}} \\
\boldsymbol{F} &\leftarrow \frac{\boldsymbol{F} \circ \sum_i \boldsymbol{G}^T \boldsymbol{P}^T (\boldsymbol{W_d} \circ \boldsymbol{I}) \boldsymbol{Q_i}^T}{\sum_i \boldsymbol{G}^T \boldsymbol{P}^T (\boldsymbol{W_d} \circ \sum_j \boldsymbol{P_j} \boldsymbol{G} \boldsymbol{F}^T \boldsymbol{Q_j}) \boldsymbol{Q_i}^T}
\end{aligned}
\qquad (10)
$$

In our implementation, we initialized the masks with random values uniformly distributed on $[0, 1]$ and after each
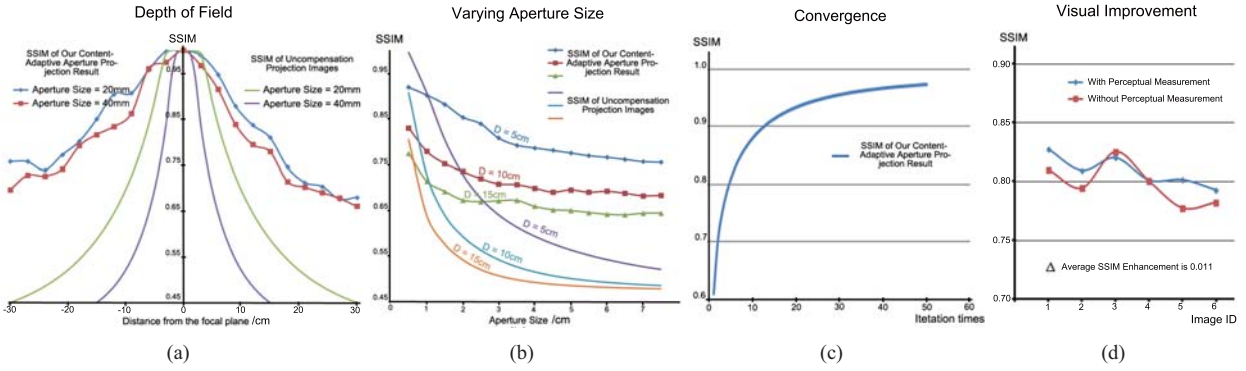
Figure 4. Quantitative evaluation of our content-adaptive aperture projection. (a) Comparisons between our defocus compensation results and the uncompensation ones under different out-of-focus distances. (b) By fixing the other parameters of the projection, our defocus compensation results and the uncompensation ones is evaluated at different aperture sizes, which is carried out at several different distances. (c) The approximation error as a function of NMF iteration times is given. (d) The improvement of the perceptual measurement incorporation is evaluated quantitatively for different target images (x-axis). In all the evaluations, structural similarity (SSIM) is used to assess the perceptual image quality.

iteration the mask elements are truncated to the range of $[0, 1]$. In conclusion, we jointly optimize both SLM layers with the proposed non-negative matrix factorization algorithm which results in a content-adaptive aperture projection.

3.4. Perceptual Measurement Incorporation

The sensitivity variations of the human vision system have been well studied. How different spatial frequencies are perceived is defined as contrast sensitivity function. Knowing which pixel is important and being focused is definitely helpful in the optimization. For this reason, perceptual measurement is incorporated into our optimization of content-adaptive aperture projection as well. We first decompose the target images and the reconstructed ones into different frequency bands by scale decomposition as used in [19][16][15]. Then according to the sensitivity of human visual system we weights each frequency band differently using the contrast sensitivity function described by [6]. With this incorporation, our objective function is adapted as

$$\underset{F, G}{\mathrm{argmin}} \| S((I - \sum_{i=1}^{n} P_i G F^T Q_i) \circ W_d) \|^2_{W_{CSF}} \quad (11)$$

where S is the scale decomposition operator. Following Mantiuk et al. [14], we approximate the nonlinear transducers modeling human contrast sensitivity with a linear matrix approximation W_{CSF}.

4. Numerical Simulation and Analysis

In order to demonstrate the feasibility of depth-of-field enhancement through our content-adaptive aperture projec-

tion, we carried out a series of numerical simulations for quantitative evaluation.

Based on the derivation and optimization in Section 3, after setting up a number of projection parameters, we perform simulations of our content-adaptive projection for several test images, including "lena", "barbara", "man", "goldhill", "baboon" and "dollar". Here, we consider the following settings: image resolution 512×512, aperture resolution 20×20, screen size $20cm \times 20cm$, focal plane distance $100cm$. In Figure 4, average results of defocus compensation were quantitatively evaluated under different screen distances and aperture sizes. Moreover, improvement of perceptual measurement incorporation is also evaluated with experiments. We use structural similarity (SSIM) to assess perceptual image quality which was introduced by Wang et al.[21] based on the degradation of structural information. Figure 4(a) displays the results of our defocus compensation results and the uncompensation ones. It shows clearly that when the screen is distant from the projector's focal plane, the projected image quality is greatly enhanced by our method. In Figure 4(b), different aperture sizes were simulated, from which we may infer that the performance gain of our method increases with an increasing aperture size. In Figure 4(c)the approximation error as a function of NMF iteration is given, and we can see the proposed algorithm converges fast. Then in Figure 4(d), results of comparative experiment based on SSIM image quality metric demonstrate the superiority of our perceptual measurement corporation, we can see that an average SSIM enhancement of 0.011 is achieved, the visual quality of the projection imagery is improved by our approach.

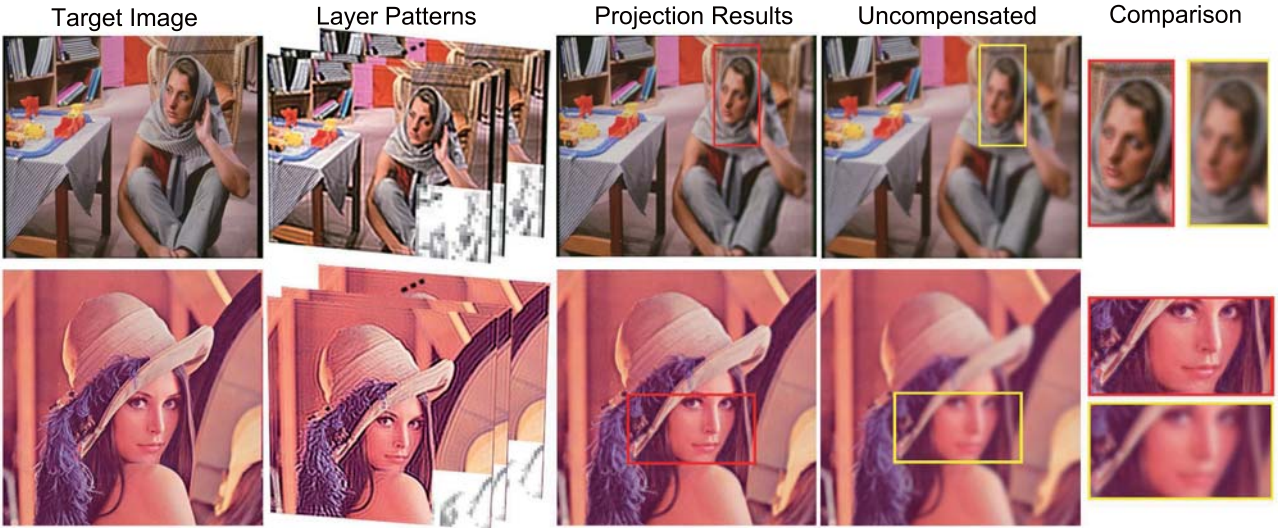Using the same parameter settings as above, Figure 5

Figure 5. Defocus compensation with high-rank coded aperture projection is shown to be effective. As shown above, we perform experiments on two different test images. From left to right: target images, frames of rear layer patterns (with aperture patterns in the subimage), photographs of prototype, uncompensated projections, and close-up comparisons. The achieved depth of field is significantly enhanced with our method.

shows two different images of our simulations that are compensated using our content-adaptive aperture projection. Each of these examples shows the target and the defocused images, as well as the compensated result. One frame of the displayed patterns for both SLMs is also shown. Comparing the uncompensated and the compensated projection, we observe that the visual quality is significantly enhanced using the proposed method. For all experiments we assume that the observer averages four subframes, which models 240 Hz SLMs for a flicker fusion threshold of 60 Hz.

For evaluation we also wish to quantitatively compare the loss in light transmission to previous work and conventional projection as well as the depth-of-field enhancement. The DOF and the light throughput of a projection system depend mainly on the effective aperture diameter and on the focal length of the objective lens. In photography, the ratio of these two parameters is generally expressed by the objectives f-stop number. A low f-stop is equivalent to a large light throughput with a small depth-of-field and vice versa. Following [18], we determine the approximate scale of the PSF in a compensated projection by finding the best match of captured projection and the original image convolved with different scales of that PSF. Using the determined scale, we compute the f-number of an objective lens with a circular aperture (and constant focal length) that would lead to the same depth-of-field ($\hat{f}/\#$). We evaluate the light throughput also in f-numbers ($\tilde{f}/\#$) with respect to a comparable standard projector equipped with an auto-iris. As shown in Figure 6, compared to conven-
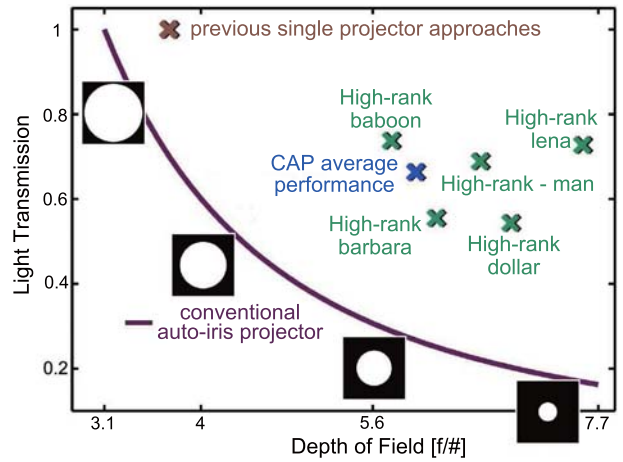

Figure 6. Loss in light transmission and DOF enhancement that our high-rank coded aperture projection can achieve is compared to previous work and conventional projection. The values are given in relative f-numbers comparing DOF ($\hat{f}/\#$) and transmission ($\tilde{f}/\#$).

tional projection and previous single projector approaches, our high-rank coded aperture projection can achieve a larger DOF without losing as much light as an auto-iris projector for which $\hat{f}/\# = \tilde{f}/\#$. Our high-rank approach also always achieve larger DOFs than the previous coded aperture projection, while maintaining a similar light transmission loss.

5. Implementation and Results

This section describes our high-rank coded aperture projection prototype; we evaluate its performance and also compare it with conventional projectors and previously proposed single-frame coded apertures.
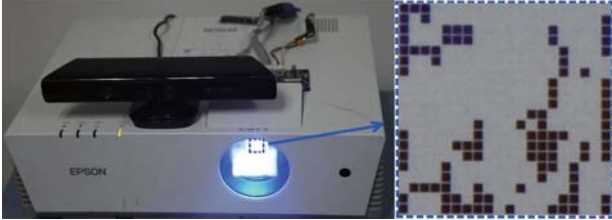
5.1. Hardware Setup

Figure 7. Prototype high-rank projector. A programmable SLM is inserted into the aperture plane of an video projector; high-rank light field factorization is employed to compute the high-speed patterns displayed on image and aperture SLM. A Microsoft Kinect captures the screen geometry.

We built a prototype projector by inserting a programmable a liquid crystal display (LCD) panel into the projector's aperture plane. The LCD is a custom designed monochrome film-compensated super-twisted nematic (FSTN) panel. Its spatial resolution is 128×32 from which about 20×20 pixels are used for the aperture SLM. The LCD has a frame rate of 60Hz; with the assumption that the observer averages four subframes, an video projection of 15Hz is built. While the achieved framerate is slightly too low for human perception, we simulate 240 Hz SLMs by taking photographs with an exposure time of 1/15 sec. That is to say, our projected light field is represented by a rank-4 light field matrix. The projector is a EPSON EMP-6110 (3 LCD). The LCD panel and the projector is driven separately via DVI links from a dual-head NVIDIA GTX550 Ti display adapter, which automatically synchronizing the display refreshes. Our prototype system is shown in Figure 7. To capture the screen geometry, we use a Microsoft Kinect.

5.2. Experiment Results

Examples of the captured projection result are shown in Figures 1, 5 and 8. In all of these experiments, we observe a significant improvement in depth of field, hence image sharpness, compared to conventional projectors. We show decomposed layer patterns (Figs 1 and 5), the depth map of the screen (Fig. 8), and comparisons to unmodified projections.

Figure 9 shows a comparison between the proposed high-rank coded aperture approach and the rank-1 coded
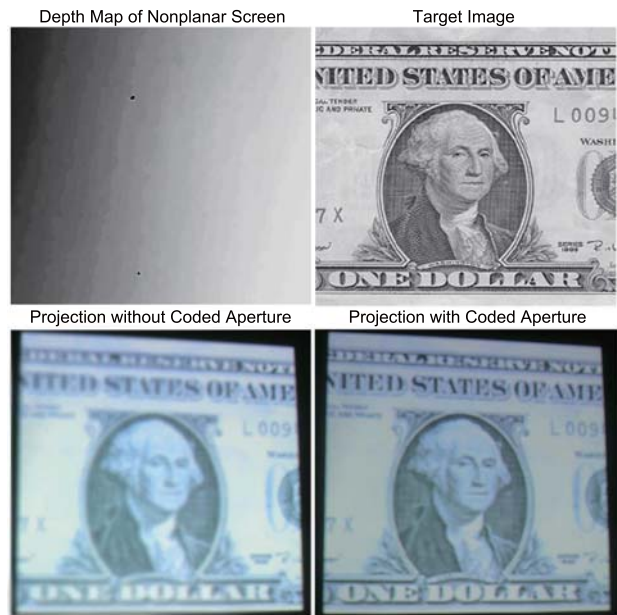

Figure 8. Another experimental result of enhanced DOF through high-rank coded aperture projection. Characters on the dollar bill are completely blurred out in conventional projection (lower left) whereas our approach preserves image sharpness (lower right). As shown in the results, our approach is a little darker due to the loss of light caused by the employed LCD. The depth map of the screen and the target image are shown in the upper row as well.
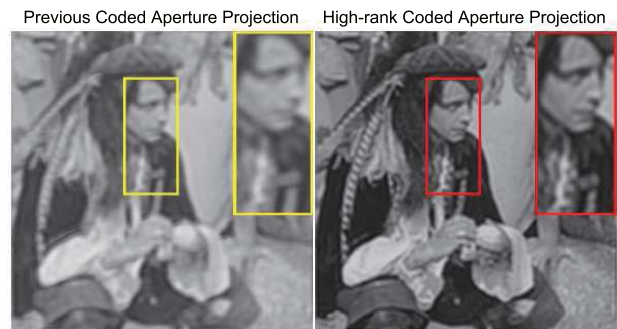

Figure 9. Compared to previously proposed coded aperture projectors [8], the proposed high-rank approach results in a significantly enhanced image quality due to the extended depth of field.

aperture projection proposed by Grosse et al. [8]. While Grosse's method is successful in compensating some of the defocus blur our high-rank methods results in a superior quality.

6. Discussion

We introduce a new type of computational projection system. Through the co-design of display hardware and

adaptive computation, we demonstrate how our high-rank coded aperture display system achieves larger depths of field than conventional and previously proposed computational projectors. Our hardware setup comprises a conventional projector with two high-speed SLMs—one in the image plane and one in the projector's aperture. The projected high-speed patterns are computed with a novel formulation for high-rank light field decomposition that exploits the limited temporal resolution and contrast sensitivity of the human visual system. The proposed display system could become the next-generation auto-iris technology for projectors, facilitating unconventional projection setups and other novel applications.

Benefits and Limitations High-rank coded aperture projectors overcome limitations of conventional and previous coded displays by significantly extending theirs depths of field while achieving a high light throughput. These capabilities enable focused projections on nonplanar screens, which is crucial for ad-hoc collaborative projection setups in everyday environments and for sophisticated projector installations in cultural heritage sites, museums, stage performances, or buildings.

The main limitations of our system are currently dictated by the employed hardware. The aperture LCD we used now has a very low light transmittance which results in a relatively high loss of light. Therefore, we trade light throughput for depth-of-field. Furthermore, the LCD has a limited contrast, exhibits slight diffraction artifacts because of the small pixel size, and only withstands the heat generated in the projector for a limited time. Furthermore, we can currently operate the SLMs in our prototype with only 60 Hz, which limits the number of patterns that the human visual system averages over to about two. Most of these limitations, such as the light loss and low contrast, however, can be addressed by using alternative SLM technologies, such as digital micromirrors or microshutters (www.pixtronix.com). While our unoptimized Matlab implementation currently takes a few seconds to compute a set of mask patterns, nonnegative matrix factorizations can be parallelized and efficiently implemented on the GPU.

Future Work In the future, we would like to incorporate more sophisticated perceptual models in our optimization framework. Furthermore, we would like to extend our work to multi-projector setups and experiment with multiple stacked SLMs in each device. We plan on exploring further applications, such as increased temporal contrast, projector de-pixelation, novel interaction paradigms, or even 4D structured illumination, that are facilitated with our computational display design.

References

[1] M. Berry, M. Browne, A. Langville, V. Pauca, and R.Plemmons. Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics and Data Analysis*, 52(1):155–173, 2007.

[2] O. Bimber and A. Emmerling. Multifocal projection: A multiprojector technique for increasing focaldepth. *IEEE TVCG*, 12(4):658–667, 2006.

[3] V. Blondel, N. Ho, and P. V. Dooren. Weighted nonnegative matrix factorization and face feature extraction. *Image and Vision Computing*, 2008.

[4] M. Brown, P. Song, and T. Cham. Image preconditioning for out-of-focus projector blur. In *Proc. CVPR*, pages 1956–1963, 2006.

[5] O. Cossairt, C. Zhou, and S. Nayar. Diffusion Coded Photography for Extended Depth of Field. *ACM Trans. Graph. (SIGGRAPH)*, 29:31:1–31:10, 2010.

[6] S. Daly. The visible differences predictor: an algorithm for the assessment of image fidelity. *Digital images and human vision*, 4:124–125, 1993.

[7] E. Dowski and W. Cathey. Extended depth of field through wave-front coding. *Applied Optics*, 34(11):1859–1866, 1995.

[8] M. Grosse, G. Wetzstein, A. Grundhofer, and O. Bimber. Coded aperture projection. *ACM Trans. Graph.*, 29(3):22, 2010.

[9] D. Lanman, M. Hirsch, Y. Kim, and R. Raskar. Content-adaptive parallax barriers: optimizing dual-layer 3D displays using low-rank light field factorization. *ACM Trans. Graph. (SIGGRAPH Asia)*, 163(6), 2010.

[10] D. Lee and H. Seung. Learning the parts of objects by non-negative matrix factorization. *Nature*, 401(6755):788–791, 1999.

[11] A. Levin, R. Fergus, F. Durand, and W. Freeman. Image and depth from a conventional camera with a coded aperture. *ACM Trans. Graph. (SIGGRAPH)*, 26(3):70, 2007.

[12] A. Levin, S. W. Hasinoff, P. Green, F. Durand, and W. T. Freeman. 4D Frequency Analysis of Computational Cameras for Depth of Field Extension. *ACM Trans. Graph. (SIGGRAPH)*, 28(3):97:1–97:14, 2009.

[13] M. Levoy and P. Hanrahan. Light field rendering. *ACM Trans. Graph. (SIGGRAPH)*, pages 31–42, 1996.

[14] R. Mantiuk, S. Daly, and L. Kerofsky. Display adaptive tone mapping. 27(3):68, 2008.

[15] R. Mantiuk, K. Kim, A. Rempel, and W. Heidrich. HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions,. *ACM Trans. Graph. (SIGGRAPH)*, 30(4):40, 2011.

[16] R. Mantiuk, K. Myszkowski, and H. Seidel. Visible difference predicator for high dynamic range images. In *IEEE International Conference on Systems, Man and Cybernetics*, volume 3, pages 2763–2769, 2004.

[17] Y. Oyamada and H. Saito. Focal pre-correction of projected image for deblurring screen image. In *Proc. CVPR*, pages 1–8, 2007.

[18] Y. Oyamada and H. Saito. Focal pre-correction of projected image for deblurring screen image. In *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, pages 1–8. IEEE, 2007.

[19] E. Simoncelli and W. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proc. ICIP*, volume 3, pages 444–447, 1995.

[20] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. *ACM Trans. Graph. (SIGGRAPH)*, 26(3):69, 2007.

[21] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Trancsactions on Image Processing*, 13(4):600–612, 2004.

[22] G. Wetzstein, D. Lanman, M. Hirsch, and R. Raskar. Tensor displays: compressive light field synthesis using multilayer displays with directional backlighting. *ACM Trans. Graph. (SIGGRAPH)*, 31(4):80, 2012.

[23] L. Zhang and S. Nayar. Projection defocus analysis for scene capture and image display. *ACM Trans. Graph. (SIGGRAPH)*, 25(3):907–915, 2006.

[24] C. Zhou and S. Nayar. What are good apertures for defocus deblurring? In *Proc. ICCP*, pages 1–8, 2009.