

Coded Aperture Projection

MAX GROSSE

Bauhaus-University Weimar

and

GORDON WETZSTEIN

University of British Columbia

and

ANSELM GRUNDHÖFER

Bauhaus-University Weimar

and

OLIVER BIMBER

Johannes Kepler University Linz

Bauhaus-University Weimar

Coding a projector's aperture plane with adaptive patterns together with inverse filtering allow the depth-of-field of projected imagery to be increased. We present two prototypes and corresponding algorithms for static and programmable apertures. We also explain how these patterns can be computed at interactive rates, by taking into account the image content and limitations of the human visual system. Applications such as projector defocus compensation, high quality projector de-pixelation, and increased temporal contrast of projected video sequences can be supported. Coded apertures are a step towards next-generation auto-iris projector lenses.

Categories and Subject Descriptors: I.3.3 [**computer graphics**]: Picture/Image Generation—*Display algorithms*; I.4.0 [**image processing and computer vision**]: General—*Image displays*

General Terms: Algorithms, Theory

Additional Key Words and Phrases: Coded Aperture Imaging, Computational Light Modulation

1. INTRODUCTION

Modern video projectors are remarkable devices that can display large imagery with high resolution, brightness, and contrast. The latest high-end models even incorporate auto-focus and auto-iris objective lenses. These can greatly enhance the temporal contrast of projected images by adjusting the aperture to the average brightness of the displayed content. Their flexibility and low cost make projectors irreplaceable for many applications including professional presentations, home entertainment, scientific visualization, as well as museum and art installations. We envision future generations of these displays as fully integrated systems with cameras, dynamically adjustable apertures and intelligent control mechanisms.

In this paper, we present solutions for taking projectors to the next level. By placing coded masks at a projector's aperture plane, we show how the depth-of-field (DOF) of a projection can be greatly enhanced.

e-mail: Max.Grosse@uni-weimar.de, wetzste1@cs.ubc.ca, bimer@uni-weimar.de, Anselm.Grundhoefer@uni-weimar.de

Permission to make digital/hard copy of all or part of this material without fee for personal or classroom use provided that the copies are not made or distributed for profit or commercial advantage, the ACM copyright/server notice, the title of the publication, and its date appear, and notice is given that copying is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers, or to redistribute to lists requires prior specific permission and/or a fee.

© 20YY ACM 0730-0301/20YY/0100-0001 \$5.00

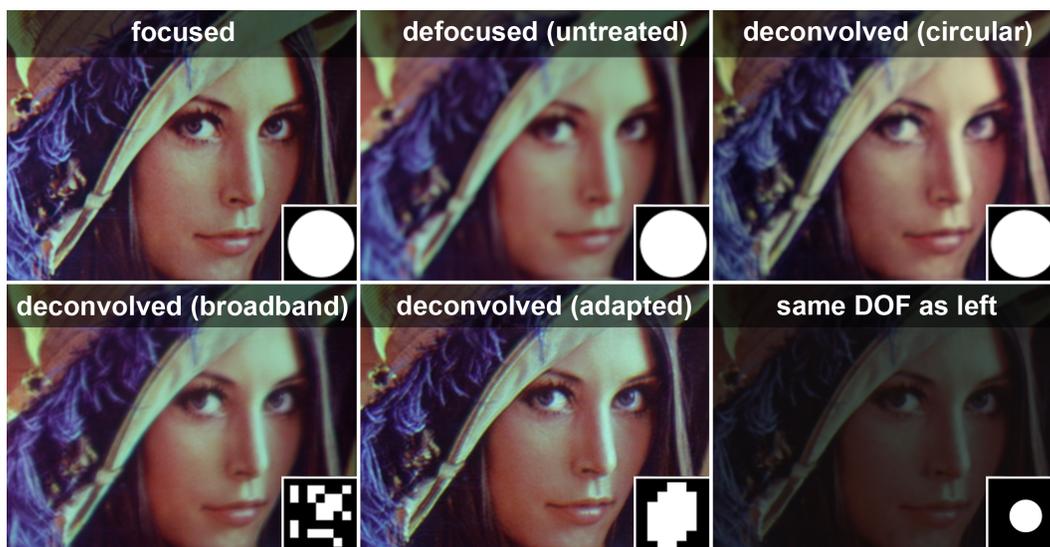


Fig. 1. Displayed imagery can be pre-compensated to enhance a projector’s depth-of-field, but the resulting quality depends on the employed aperture shape. A projected image in focus (upper left), and with the same optical defocus (screen located at 2m distance to focal plane) modified in several different ways: with circular aperture - untreated and deconvolved with corresponding PSF, compensated with static broadband coded aperture and with dynamic coded aperture. The result of the dynamic aperture is optimized for a viewing distance of at least 50cm and for an image diameter of 10cm where no artifacts should be visible. The deconvolved images are computed from the corresponding apertures (shown in the sub-figures). While all other images are adjusted to a similar brightness to enable a better comparison of focus, the lower right image is captured using the same exposure as the image to its left. A small circular aperture has been applied that achieves the same depth-of-field as with the adaptive coded aperture at the cost of a significant loss of brightness. Note that the first four and the last two images were displayed with two different projector prototypes (see figure 4), which have slightly different color gamuts and intensity ranges.

This allows focused imagery to be shown on complex screens with varying distances to the projector’s focal plane, such as projection domes or cylindrical canvases. We demonstrate that integrated broadband masks and adaptive dynamic apertures outperform previous methods of defocus compensation for circular aperture stops. In addition, our dynamic apertures can perform the type of contrast enhancement employed by common auto-iris projection lenses. Furthermore, they enable high-quality de-pixelated projections that are beneficial for rear-projection TV sets and other close-view displays.

Specifically, we make the following contributions:

- the introduction of coded apertures to display devices
- an integrated hardware solution and corresponding algorithms to improve the depth-of-field and temporal contrast of projectors
- a novel, content-dependent algorithm for computing dynamic adaptive apertures that is based on the human visual system

2. RELATED WORK

Defocus Compensation for Projectors. Enhanced depth-of-field projection with multiple displays has been demonstrated by Bimber and Emmerling [2006]. The spatially varying point spread functions (PSFs)

are measured for each projector and used to determine the contributions of multiple overlapping projections with differently adjusted focal planes that lead to the least amount of defocus.

Single projector techniques generally pre-sharpen displayed imagery digitally, so that an optically defocused projection appears close to the original, un-blurred image. Brown et al. [2006] and Oyamada and Saito [2007] proposed to compensate the defocus by performing Wiener deconvolution in the frequency domain. Zhang and Nayar [2006] formulated the image correction as a constrained optimization problem in the spatial domain using a convolution matrix. The constraints ensure that the solution is within the dynamic range of the projector. All of these techniques measure the PSFs using projected code patterns except for [Oyamada and Saito 2007], where it is estimated using pattern matching.

Multi-focal projection with several displays may be impractical due to the required number of projection units. All single projector approaches, on the other hand, share two limitations. First, they do not run at interactive frame-rates, which prevents them from displaying dynamic content. Second, the amount of defocus that can be compensated through digital processing alone is clearly limited. Because of a circular aperture stop and a resulting PSF that contains many zero frequencies, high image frequencies are irreversibly lost in out-of-focus regions. A pre-compensation of the optically low-pass filtered projected light often leads to dominant ringing artifacts, especially with direct inverse methods such as Wiener deconvolution.

An extended depth-of-field projector was introduced by Dowski and Cathey [1998]. The authors proposed to insert a cubically shaped lens element (phase plate) into the projected light path, which results in a projector PSF that is independent of the distance to the focal plane (see [Dowski and Cathey 1995]). For compensation, projected images are deconvolved with the spatially uniform PSF. While this design allows high light-throughput, it has clear limitations. The precise manufacturing and insertion of the lens element is costly and difficult. Once integrated into the display, it cannot be 'switched off'. Even image regions that would be in-focus without the phase plate are blurry. By using coded apertures instead, we do not alter the image sharpness for in-focus areas, but only improve it for out-of-focus regions. Furthermore, a dynamic aperture adds the benefits of higher temporal contrast just like auto-iris projector lenses do, and can be turned off so as not to affect the projection at all. Levin et al. [2009] recently analyzed design parameters for lens elements that optimize the depth-independence of the corresponding PSF in cameras. Applying these paradigms to projectors has the same disadvantages as the extended depth-of-field projector proposed by Dowski and Cathey [1998].

Coded Aperture Imaging. Joint optical light modulation and computational reconstruction for enhanced image capture has many applications. Among these are extended depth-of-field photography [Dowski and Cathey 1995], split field-of-view image capture [Zomet and Nayar 2006], motion deblurring [Raskar et al. 2006], light-field capture for synthetic image refocus [Liang et al. 2008; Bando et al. 2008; Veeraraghavan et al. 2007; Ng 2005], and depth reconstruction [Levin et al. 2007; Bando et al. 2008; Moreno-Noguer et al. 2007]. An analysis of different aperture patterns for high-quality defocus deblurring was presented by Zhou and Nayar [2009]. Their codes are optimized for specific noise parameters of cameras, but this does not apply to projectors, as there is no noise contribution.

To our knowledge, we are the first to apply coded apertures to displays.

3. CODED APERTURE PROJECTION PRINCIPLE

Common projection displays have large aperture stops to optimize light transmission, and therefore narrow depths-of-field. When displaying imagery onto non-planar canvases or when fronto-parallel display positions are not feasible, projected content will be blurred due to optical defocus. This is often described as a convolution of the original image with a filter kernel that corresponds to the aperture of the imaging device (PSF, figure 2 left). The scale of the kernel is directly proportional to the degree of defocus:

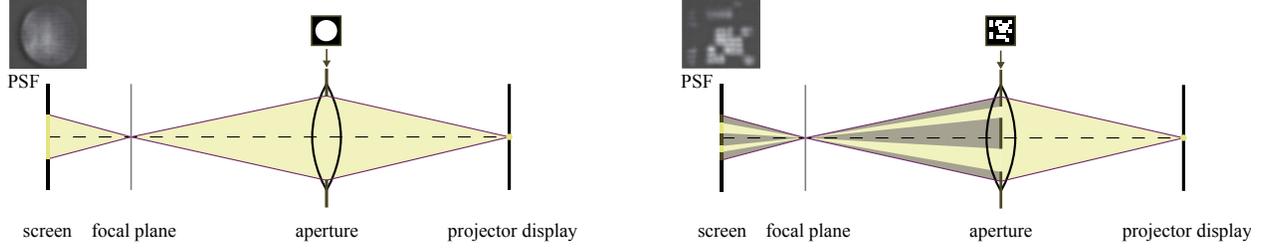


Fig. 2. Different aperture patterns lead to different point spread functions. A defocused projection of a point results in an image of the utilized aperture, as shown in case of a circular aperture (left) and a coded aperture (right).

$$i_p = k_s \otimes i_d, \quad (1)$$

where i_d is the displayed image, k_s the aperture kernel at scale s , and i_p the optically blurred projection. Deconvolution, i.e., convolving an image with the inverse aperture kernel, will digitally sharpen an image and consequently compensate optical defocus:

$$i_d = k_s^{-1} \otimes i. \quad (2)$$

Here k_s^{-1} is the inverse aperture kernel, and i the original desired image. Convolution and deconvolution can be modeled easier in the frequency domain, rather than in the spatial domain, where a convolution corresponds to a multiplication $I_p = K_s \cdot I_d$ and deconvolution is equivalent to a division $I_d = I/K_s$. I , I_d , I_p , and K_s are the Fourier transforms of i , i_d , i_p , and k_s respectively. Deconvolution is an ill-posed problem and performing it directly through division in frequency space is generally avoided. Zeros in the Fourier representation of the convolution kernel and image noise often lead to ringing artifacts. For projector defocus compensation, however, the images are not measurements, but noiseless digital footage. This makes deconvolution by direct inverse filtering feasible if the kernel is broad-band, i.e. has no zeros in its Fourier transform. Instead of Wiener deconvolution, we propose to apply a regularized inverse filter to compute a displayed compensation image I_d as

$$I_d = \frac{K_s^*(f_x, f_y)}{|K_s(f_x, f_y)|^2 + \alpha |L(f_x, f_y)|^2} I(f_x, f_y), \quad (3)$$

where $L(f_x, f_y)$ is the Fourier transform of the discrete Laplacian operator and $\alpha = 0.01$ is a user-defined smoothing parameter. The regularization term $\alpha |L(f_x, f_y)|^2$ prevents ringing artifacts by smoothing out any possible division by zero for higher frequencies. If chosen too high, however, the regularizer prevents any significant sharpening. Intuitively, it can be understood as an approximation to the unknown signal-to-noise ratio that is employed in the same fashion for Wiener deconvolution.

Once the deconvolution has been computed in the frequency domain, the result I_d is inverse Fourier transformed and projected as a compensation image. Note that different image portions are deconvolved with kernels of different scales s . The scales correspond to the amount of measured defocus at the corresponding screen areas (see [Bimber and Emmerling 2006]). Thus, the larger the desired depth-of-field, the more scales are necessary.

The main limitation of previous projector defocus compensation approaches are ringing artifacts in the pre-filtered images. These are mainly due to the aforementioned divisions by zeros introduced by circular projector apertures. The resulting PSFs act as low pass filters on projected light by irreversibly canceling

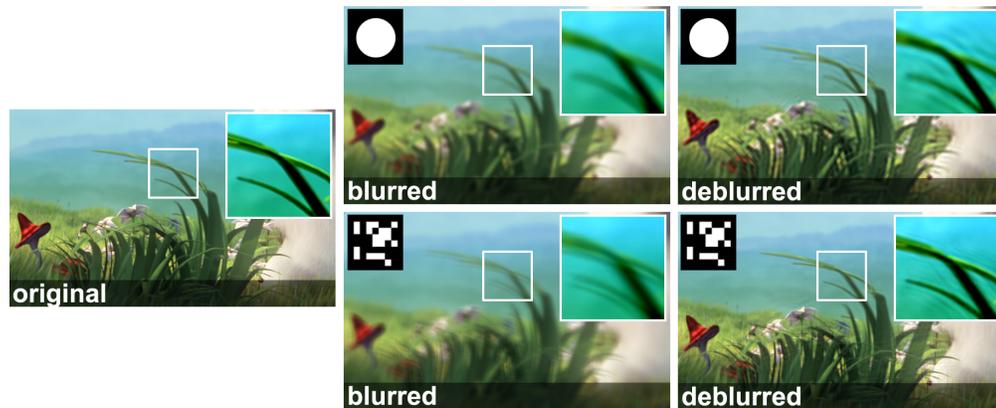


Fig. 3. Deconvolution with different kernels. Blurring an image (left) can, depending on the kernel, filter out important frequencies in the blurred result (center column), which often results in ringing artifacts after a deconvolution (upper right). A broad-band kernel can significantly improve the quality of the deblurred result (lower right). The close-up regions are contrast enhanced for improved visibility.

out high spatial frequencies. The loss of high spatial frequencies and resulting ringing in the compensation images is depicted in figure 3, upper row. Decreasing the size of the aperture opening reduces the number of low Fourier magnitudes, thus it effectively enhances the depth-of-field of a projection system. Using narrower aperture openings (up to pinhole size), however, will naturally decrease the light throughput significantly, which is unacceptable for most projection-based displays.

In the remainder of this paper, we will explore the use of coded apertures that optimize a projector’s depth-of-field without the need of additional refractive lens elements. We show how static [Grosse and Bimber 2008] and dynamically adjusted attenuation masks in the projector’s aperture plane can preserve sharp image features with much fewer ringing artifacts (see figure 3, lower row) for defocused projections while preserving a high light transmission. The employed apertures are either designed to be broadband in frequency domain or optimized to the frequency band of the current image. They comprise higher Fourier magnitudes than circular apertures with the same light throughput.

Previously proposed single projector defocus compensation approaches apply either Wiener filtering [Brown et al. 2006; Oyamada and Saito 2007] or a least-squared error solution in the spatial domain [Zhang and Nayar 2006]. While the former approach is very similar to our regularized inverse filter (equation 3), it has not been shown in those papers that a solution can be computed in real-time. Section 5 of our submission is mainly dedicated to the necessary data-preprocessing and our real-time GPU implementation. Hence, we apply a similar algorithm but show how to make it work in real-time.

4. PROTOTYPE DESIGNS

In order to demonstrate the feasibility of depth-of-field enhancement through coded projector apertures we built two prototypes. Our first prototype utilizes a static attenuation mask. We integrated a programmable liquid crystal display into a second prototype to experiment with adaptive aperture patterns.

4.1 Static Broadband Aperture

In our first prototype, we inserted a static attenuation mask into the aperture plane of the objective lens, as shown in figure 4 (left). The projector is a Sony VPL-CX80 XGA 3-chip LCD projector and the applied aperture pattern was found by Veeraraghavan et al. [2007] by seeking a 7x7 binary broadband mask that



Fig. 4. Two prototypes: a static attenuation mask integrated into the aperture plane of a projector improves digital defocus compensation through inverse filtering (left). Replacing the aperture with a transparent liquid crystal array allows dynamic attenuation mask patterns to be encoded (right). We compute the adaptive patterns by optimizing for light throughput while preserving spatial frequencies that are perceivable by a human observer. Both approaches significantly improve depth-of-field through inverse light filtering when compared to circular aperture stops. While static coded apertures are easier to realize and less expensive, adaptive coded apertures are more effective.

maximizes the Fourier-magnitudes of the zero padded convolution kernel.

Although photographic film is an obvious option for producing high-contrast aperture patterns, our initial experiments showed that it does not withstand the heat generated in the projector. Hence, we printed our codes on transparencies, which proved more stable under varying temperatures. In a commercial display the masks could be manufactured from any heat resisting material, thus providing an even higher contrast. They represent a low-cost solution for coded aperture projection.

4.2 Programmable Aperture

We implemented adaptive coded apertures by integrating a programmable liquid crystal array (LCA) into the projector’s aperture plane, as illustrated in figure 4 (right). The LCA is an Electronic Assembly EA DOGM132W-5 display with a resolution of 132x32, 20Hz frame rate, and $0.4 \times 0.35 \text{ mm}^2$ pixel size. It is controlled via USB with an Atmel ATmega88 microcontroller. The projector is a BenQ 7765PA XGA (DLP).

5. DEFOCUS COMPENSATION WITH STATIC CODED APERTURES

In the following subsections we will discuss implementation details for depth-of-field enhancement through inverse light filtering with static coded apertures. All of our computations are performed on programmable graphics hardware and run at interactive frame rates. We will discuss the necessary pre-processing and data acquisition steps in section 5.1, the computations that need to be performed at each frame in section 5.2, and present results in section 5.3.

5.1 Defocus Estimation and Data Preprocessing

Our defocus compensation approach relies on knowledge of the amount of defocus at each projector pixel, which directly corresponds to the scale of the convolution kernel or PSF at this point. We measure the spatially varying defocus by projecting structured light patterns and capturing them with a camera as proposed by Bimber and Emmerling [2006]. This is a one-time calibration step and takes less than a minute. Any other defocus estimator could be applied as well. Given the kernel scales for each pixel, we compute a non-uniform sub-division that partitions the projection images into regions with similar amounts of defocus

and, consequently, kernel scales. This image sub-division is constant for static settings and is therefore pre-computed along with the properly scaled inverse PSFs.

As outlined in section 3, a defocus compensation can be performed by image deconvolution with an appropriately scaled PSF. Two different approaches were proposed for this purpose: Brown et al. [2006] performed the deconvolution for each kernel scale on the whole image, whereas Oyamada and Saito [2007] subdivided the image into a uniform grid of sub-image patches prior to the actual compensation. For the latter approach, only the local PSF scales that affect a patch need to be taken into account. Local deconvolution on uniform patches is more efficient, but it does not consider the (possibly non-uniform) local distribution of blur levels, and can therefore still cause a significant amount of overhead. To overcome this, we propose a hybrid approach for image subdivision that combines the measured non-uniform defocus map and a uniform patch structure based on optimal runtime performance.

Despite the well known Fast Fourier Transform (FFT) runtime of $O(N \log N)$, for power-of-two sized images with N pixels, we found that, in practice, the performance of hardware and software specific FFT implementations can vary significantly. Figure 5 (left) shows our measurements of the CUDA FFT performance on a NVIDIA GeForce 8800 Ultra graphics board. For all possible input sizes that are a power of two, the time per pixel required when performing a FFT was measured. Patch sizes of 1024×256 and 1024×1024 are most efficiently processed in this example.

Given the measured defocus map and the FFT performance of all patch size combinations, we perform a top-down quad-tree image subdivision up to a predefined level. The subdivision is started from the nearest higher power-of-two image size to ensure that every node also has a power-of-two-size. The tree is then traversed bottom-up, where we successively merge patches at each level if this leads to more efficient results. A merge operation is only performed among child nodes of the same parent. The efficiency for each possible merge is given by the number of different defocus levels in the resulting node multiplied with the node's FFT computing time. Figure 5 shows an exemplary patch subdivision.

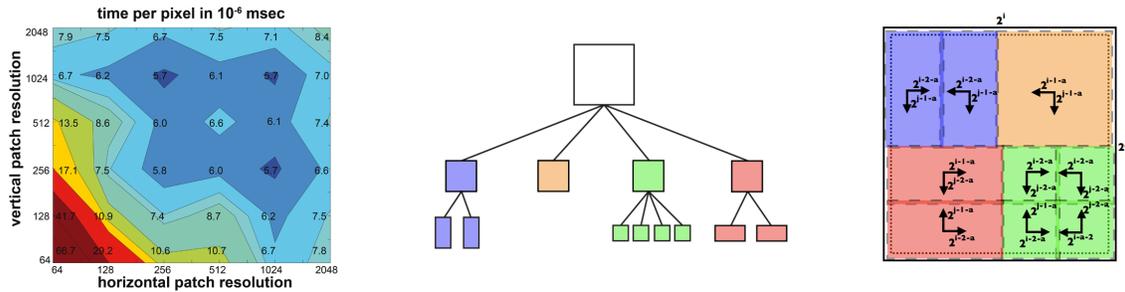


Fig. 5. Measured efficiency (average processing time per pixel) for each power-of-two patch size on a NVIDIA GeForce 8800 Ultra graphics board (left), example of a partially expanded and collapsed quad-tree (center) that corresponds to the subdivision (right) which illustrates merge possibilities and overlap shifts.

All patches have to overlap adequately to provide sufficiently large blending regions. To ensure this, we pad the original image with zeros before up-sampling it to the proper power-of-two resolution. This padding area is $2^{x-a} \times 2^{y-a}$, where 2^x and 2^y are the edge lengths of the up-sampled image. Then, each patch is shifted towards its split edges as indicated in figure 5 (right). The amount of a shift depends on the patch size: an edge of length 2^i is shifted by 2^{i-a} . In our implementation, a equals 3. This operation always ensures a sufficient overlap for blending regions in all levels.

All these steps are carried out off-line. The next section explains the run-time computations that are performed for each frame.

5.2 Real-Time Compensation on the GPU

In each frame, the general computation steps are: up-sample the image to the nearest power-of-two, convert to luminance, apply the pre-computed patch subdivision, compute the Fourier transform of the patches, apply an inverse filter by dividing through the appropriate PSF, transform the patches back to the spatial domain, blend them, and add the unprocessed chrominance terms back in. Each step will be discussed in the following.

As mentioned in section 5.1, the original image is up-sampled to the nearest power-of-two resolution to achieve the lowest possible computation times. Up-sampling also allows more integer sized kernel scales to be used. As the human visual system is much more sensitive to luminance than to chrominance variations, we apply the deconvolution only to the luminance channel, instead of performing the compensation on each color channel separately. Afterwards, the compensated luminance channel and the original chrominance data are merged and transformed back to RGB space before displaying the image.

For the compensation step, we subdivide the upscaled luminance image into the pre-computed patches (see Sec. 5.1). Because it is not possible using the CUFFT API to perform a FFT only on a subregion of an image, all patches have to be reordered into a single one-dimensional array. This array has the same number of elements as the original image, thus no overhead is created. Consequently, a FFT of a consecutive subset of that array is the transform of the according patch. After Fourier transforming all patches, a deconvolution for each required scale of all patches as described in equation 3 is carried out. Deconvolution requires one or more multiplications of each patch with the corresponding inverse kernels, depending on the amount of different scales required for that patch. The result is transformed back into the spatial domain, linearly blended in the overlapping regions, down-sampled, and recombined with the original chrominances to get the final compensation image.

The entire workflow is illustrated in figure 6. The patches in this example have three, one, three, and two different PSF scales, respectively. Therefore, equation 3 as well as the subsequent IFFTs have to be executed nine times (for all required scales of each patch), while the FFT is computed four times (once per patch).

5.3 Static Aperture Results

Figure 7 compares the amount of low Fourier magnitudes of the chosen binary broad-band pattern [Veeraraghavan et al. 2007] and a circular mask with the same light transmission for different PSF scales. The de facto opening for both apertures is the same. For this plot, the scaled patterns are zero-padded to original image resolution, and show the amount of resulting Fourier magnitudes below 7% of the maximum value. The coded aperture outperforms the circular mask in all scales.

With our current prototype, we achieve 10-17 *fps* for XGA resolution and 1-14 scales (see figure 7). Upcoming graphics hardware has the potential to enable real-time performance for a large number of scales. Figures 1 and 8 show examples of depth-of-field enhanced projection using the static mask. As seen in figure 8, the optical defocus caused by a projection onto a spherical screen can be compensated by our approach. Further evaluation of static apertures is presented in the course of a comparison with adaptive apertures in section 7.

One limitation of statically coded apertures is that they reduce the light throughput by a constant factor which is independent of the actual image content. Adapting the aperture dynamically with respect to the input allows us to optimize light throughput and deconvolution quality for each projected frame as described in the following section.

6. DEFOCUS COMPENSATION WITH ADAPTIVE CODED APERTURES

As explained above, increasing the depth-of-field with a static coded broadband aperture comes at the cost of constantly decreased light transmission, which is one of the most crucial aspects of all projector-based

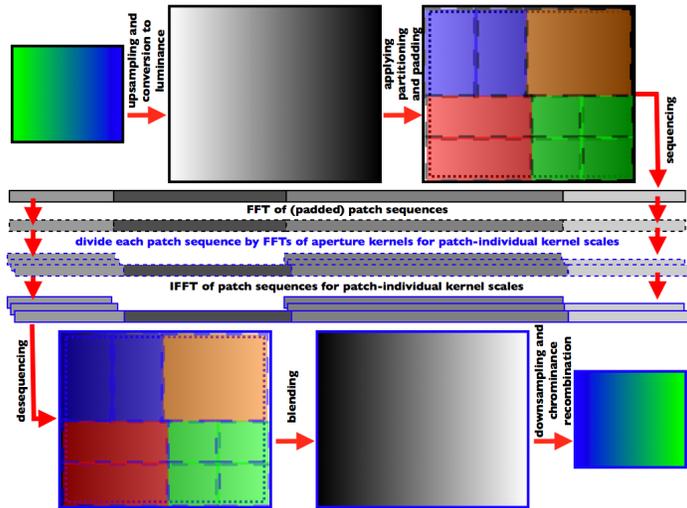


Fig. 6. A workflow diagram for per-frame defocus compensation on the GPU (from top-left to bottom-right): luminance conversion and up-sampling, image partitioning, patch sequencing, FFT of patch sequences, divisions in frequency domain for all scales required for each patch, IFFT for all patches and all scales, patch de-sequencing, blending of overlapping patches in the spatial domain, down-sampling to original size and re-combination with the original chrominances.

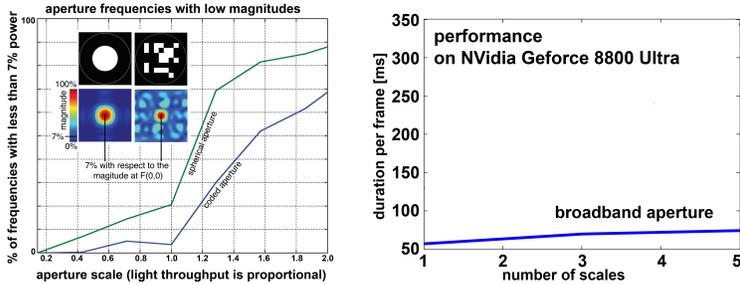


Fig. 7. Low Fourier magnitudes cause ringing after inverse filtering. The number of low frequencies of a broadband code is always less than that of a Gaussian PSF (left). This causes less ringing artifacts even at higher scales. The number of required scales, i.e. the amount of different blur levels on the screen, is linearly proportional to the processing time of defocus compensation (right).

display systems.

In this section we present a generalization of the techniques presented in section 5 to dynamic adaptive aperture patterns. The data pre-processing including image segmentation and compensation via deconvolution can be performed as explained in the previous section. We now show how dynamic apertures can be computed based on an analysis of the projected image content. This analysis employs an intuitive model of the human visual system (HVS) and allows us to determine and filter out spatial frequencies of the original image that cannot be perceived by a human observer (Sec. 6.1). An adaptive aperture can then be computed by maximizing its light transmission while preserving the perceivable frequencies, rather than being restricted to support a constant but broad frequency band (Sec. 6.2). Sections 6.3 to 6.5 discuss how temporal consistency of the computed apertures is achieved and how physical constraints are incorporated.



Fig. 8. Projection onto a hemispherical dome (5 different blur levels, depth of 40 cm) focused at the outer frame for demonstrating the depth-of-field enhancement (left). Defocus with spherical aperture at inner close-up section (center), and focus improvements with our static aperture mask (right).

In section 6.6 we show that our adaptive dynamic apertures produce better results than previous methods with the same or even an increased amount of light transmission.

6.1 Image Analysis

Although the HVS has enormous capacities such as dynamic range, spatial resolution and sensitivity to different wavelengths, it also has limitations. As all lenses in optical systems, the eyeball has a modulation transfer function (MTF) that acts as a low-pass filter on the incident light, thus attenuating or even removing high spatial frequencies. Furthermore, the sensory receptor cells in the retina are not infinitesimally small. Thus, the spatial frequencies that can be resolved on the retina are limited by the sampling theorem. We present a simple model of these optical limitations of the HVS that guide the computation of content-adapted dynamic apertures.

The sensitivity variations of the HVS according to spatial frequencies f_x, f_y are well studied and mathematically defined by the contrast sensitivity function (CSF) $S_{csf}(f_x, f_y)$. Various definitions of this function appear in the literature; we use the one described by Daly [1993]. The CSF depends on the viewing conditions only, not on the actual content. The sensitivity is defined as the inverse of the contrast required to produce a threshold response $S_{csf}(f_x, f_y) = 1/C_{thresh}(f_x, f_y)$, with C_{thresh} being the threshold contrast. Using the definition of Michelson contrast, this is given as $C_{thresh}(f_x, f_y) = \Delta L(f_x, f_y)/L_{mean}$, where ΔL is the necessary luminance difference given in cd/m^2 and L_{mean} is the mean image luminance. An absolute luminance threshold map can be computed as:

$$\Delta L(f_x, f_y) = \frac{L_{mean}}{S_{csf}(f_x, f_y)} \quad (4)$$

The sensitivity for frequencies in the range of 2 to 4 cycles per visual degree (cpd) is highest; it drops for lower and higher frequencies. As outlined by Ramasubramanian et al. [1999], an overestimation of the threshold ΔL for lower frequencies can be avoided by setting the frequency sensitivity below 4 cpd to its maximum. The threshold map is shown in figure 9.

The data for equation 4 was acquired by presenting human subjects gratings of various frequencies and intensities on a constant background luminance [Daly 1993]. The equation is a mathematical model that allows a single spatial frequency to be altered below the allowed threshold without being perceived. The actual threshold represents a 50% probability for the change to be detected by a human observer.

For computing our dynamic apertures, we wish to eliminate all frequencies that do not contribute to perceivable image fidelity. An intuitive understanding of this is reached by regarding the CSF as the equivalent

noise of the HVS. Similar approaches have been used to add noise with a spectral shape of the CSF to images [Ahumada and Watson 1985]. Note that we do not want to add noise, but rather remove frequencies from the original footage that are within the range of the HVS noise. Image compression techniques such as the JPEG standard [ITU 1993] employ similar ideas by quantizing the image frequencies. High frequencies with low magnitudes are usually quantized to zero.

The net-effect of modifying more than one spatial frequency within the threshold given by equation 4 can be expressed using probability summation as $P_{net} = 1 - \prod_{i \in f_x, f_y} (1 - P_i)$ ([Daly 1993]). It is easy to see that modulating multiple frequencies, even within the threshold given by equation 4, results in a noticeable effect. We compensate for the mutual interaction between modulated frequencies by lowering the threshold for each frequency to a detection probability that is very close to zero, so that any combination of altered frequency has still a low probability of being noticeable¹.

The Fourier magnitudes of an image converted to absolute luminance values $L(f_x, f_y)$ correspond to the amount of spatial frequencies in the image. With this information, we can calculate a binary importance mask for the image frequencies as:

$$M(f_x, f_y) = \begin{cases} 1, & |L(f_x, f_y)| \geq s\Delta L(f_x, f_y) \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

As seen in figure 9, filtering the Fourier transform of an image with the binary importance mask M allows us to remove spatial frequencies that do not modify the perceived image content for specific viewing conditions (i.e. a fixed adaptation luminance, viewer position, and screen size).

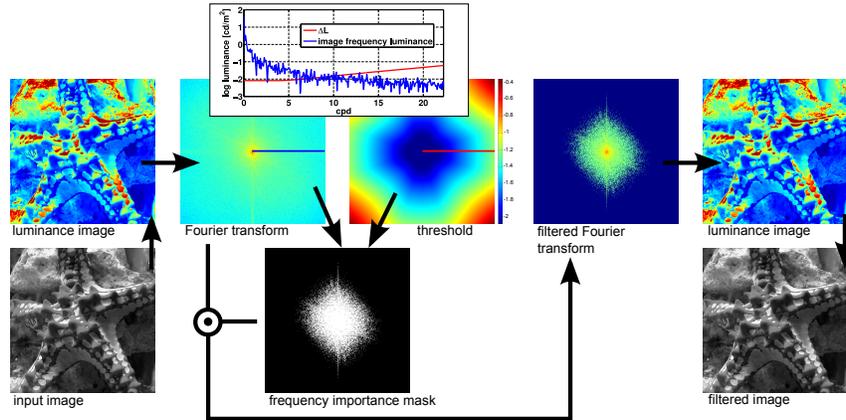


Fig. 9. Adaptive thresholding: The original image is converted to absolute luminance values. A binary frequency importance mask can be computed by thresholding the image frequencies according to a model of the HVS. The difference between the original and filtered images is not perceivable under specific viewing conditions. Scanlines of the image’s Fourier transform and the threshold map $\Delta L(f_x, f_y)$ are shown in the plot.

6.2 Dynamic Aperture Adaptation

Next, we will describe how to compute the dynamic aperture itself. We define the aperture as the sum of its individual pixels $a(x, y) = \sum_{i=1}^N a_i p_i$, where p_i is the pixel p at x_i and y_i (with a total of N pixels) and

¹In practice we scale the threshold luminance in equation 4 by a user defined parameter s . A scaling factor of 0.01 has proven successful for our experiments.

$a_i \in [0, 1]$ is its transmissivity. The Fourier transform of the aperture is $F\{a(x, y)\} = A(f_x, f_y) = \sum_{i=1}^N a_i P_i$. Our dynamic apertures should support all important frequencies in the input image with a minimal variance of their Fourier transform. In addition, they should maximize light throughput.

The variance of the aperture’s modulation transfer function (MTF) is a measure for how different frequencies are attenuated. Minimizing it for all important frequencies ensures that they are all supported. A similar criterion was employed by Raskar et al. [2006] for a one-dimensional binary temporal mask. The minimization can be mathematically expressed as an optimization problem:

$$\underset{a}{\text{minimize}} \|\mathbf{M}\mathbf{B}a - b\|_2^2, \quad (6)$$

where b is a vector containing only 1s and $a_i > 0$ are the aperture pixel intensities. We do not enforce the pixel intensities to be below 1 in this formulation, but simply scale the resulting values so that the maximum is 1. This is equivalent to scaling the MTF and does not affect the variance criterion. \mathbf{M} is a diagonal matrix containing the binary frequency importance mask values described in section 6.1. \mathbf{B} is a matrix with orthogonal basis functions in its columns which represent the optical transfer function (OTF) of the N individual aperture pixels P_i . This results in a linear system of the form $\mathbf{M}\mathbf{B}a = b$. Solving this heavily over-determined system in a least-squared error sense with the additional constraint to minimize $\|a\|_2^2$ will minimize the variance of the Fourier transform of the aperture for important frequencies. This formulation also intrinsically maximizes the light transmittance of the resulting aperture, because a small squared 2-norm of a ($a_i \geq 0$) also minimizes the variance of the normalized pixel intensities in the spatial domain.

The linear system can certainly be solved with standard approaches, such as the conjugate gradient method for the normal equations or non-negative least squares solutions. However, this would not allow sufficiently high frame rates on commonly available computer hardware for standard image resolutions of 1024×768 and higher. Thus, we propose to solve the system using the pseudo-inverse matrix, an approach that has previously been employed in solving inverse problems for light projection in real-time [Wetzstein and Bimber 2007]. Computing solutions of linear problems using the pseudo-inverse minimizes the least-squared error and the 2-norm of the resulting vector, thus solving the variance and the light transmittance problem at the same time.

Reformulating our problem results in $a = \mathbf{B}^+ \mathbf{M}^+ b$, where $^+$ denotes the pseudo-inverse matrix. Since \mathbf{M} is a binary diagonal matrix then $\mathbf{M}^+ = \mathbf{M}$. \mathbf{B} comprises the set of orthogonal Fourier basis functions as its columns, thus $\mathbf{B}^+ = \mathbf{B}^*$. We need to employ the conjugate transpose \mathbf{B}^* , because \mathbf{B} is complex, hence:

$$a = \mathbf{B}^* \mathbf{M} b \quad (7)$$

In this formulation \mathbf{B}^* can be easily pre-computed. During run-time we solve the system with a matrix-vector multiplication. Since the solution a can contain negative values we clip these values and scale the result so that the maximum value is 1.

6.3 Enforcing Temporal Consistency

Altering the aperture of the projector during image display affects the average luminance from frame to frame. In order to avoid noticeable intensity variations in form of flickering during subsequent frames with a similar average luminance, we employ the model for temporal luminance adaptation of the HVS that was proposed by Durand and Dorsey [2000]. This was previously used for interactive tone mapping [Durand and Dorsey 2000; Krawczyk et al. 2005] and adaptive radiometric compensation [Grundhofer and Bimber 2008].

The luminance adaptation process can be described as an exponential decay function [Durand and Dorsey 2000]:

$$L_t = L_{t-1} + (L_t^* - L_{t-1}) \left(1 - e^{-\frac{t}{\tau}}\right), \quad (8)$$

where T is the time step between consecutive frames in seconds, τ is a time constant, we assume photopic vision, thus $\tau = 0.1$ sec. L_t is the smoothed average aperture luminance at time step t and L_t^* the unsmoothed average luminance of the normalized dynamic aperture.

Thus, instead of using the highest possible mean transmittance for a dynamic aperture computed at frame t , we use the same aperture scaled so that its mean transmittance does not, according to the model of luminance adaptation, result in perceivable flickering compared to the last frame.

6.4 Accounting for Different Amounts of Defocus

Although equation 7 presents a valid solution for computing an adaptive, content-dependent aperture, this only accounts for a single aperture scale, which corresponds to exactly one fronto-parallel plane with a fixed distance from the optical focal plane. The distance is proportional to the scale of the aperture in the projected image.

In order to account for multiple scales we propose to choose the largest measured defocus and as a result the largest required scale for computing the dynamic aperture. Recall that a decreased scale in the spatial domain is equivalent to an increased scale in the frequency domain. Consequently, if the Fourier transform of the dynamic aperture with the largest scale supports all important frequencies, so will all the smaller spatial scales. We assume that displayed images follow natural image statistics, which means that their spectrum somewhat resembles the shape $1/f$ in the frequency domain. It is important to note that natural image statistics determine the general shape of our binary frequency importance mask M (equation 5). Assuming that lower frequencies have higher magnitudes leads to an M which is always completely filled. Since the dynamic aperture is optimized to support all frequencies within M for the largest spatial scale, all smaller scales (i.e., less-defocused areas) will contain M as a subset because they support a larger area in the frequency domain. Therefore, we compute an adaptive aperture using equation 7 for the largest measured amount of defocus and resample the result to the resolution of the physical aperture display using integral sampling.

6.5 Incorporating Physical Constraints of the Prototype

The liquid crystal arrays in our prototypes are currently limited to binary mask patterns. The resulting values of equations 6 and 7 are continuous, and cannot be displayed directly on binary LCAs. In order to solve equation 6 for a binary LCA, discrete optimization approaches, such as integer programming, would have to be employed. However, these would not account for the limited contrast of LCAs. Furthermore, discrete optimization approaches are generally much slower than continuous methods. In order to achieve interactive frame rates for dynamic image content, we first calculate a continuous adaptive aperture mask with equation 7 on the GPU, and then apply a simple continuous-to-binary mapping scheme. Given the minimum and maximum aperture display transmittance (τ_{min} and τ_{max}) we compute the binary aperture a_{bin} as follows:

$$a_{bin}(x, y) = \begin{cases} 1, & a(x, y) \geq \tau_{min} + \frac{(\tau_{max} - \tau_{min})}{\delta} \\ 0, & otherwise \end{cases} \quad (9)$$

The parameter $\delta \geq 1$ represents a tradeoff between light transmissivity of the binarized aperture and accuracy. The minimum and maximum display transmittances are given by τ_{min} and τ_{max} respectively. Empirically, we found $\delta = 2$ to be appropriate.

6.6 Adaptive Coded Aperture Results

We implemented the computation of our dynamic aperture patterns entirely on the GPU. Therefore, the B^* matrix is precomputed and uploaded onto the graphics hardware memory. We store the matrix as complex

floating point values and limit the matrix size to $1024^2 \times 7^2$ entries, which requires 383 MB of memory storage. Hence, the compensation is performed on images with a resolution of 1024×1024 and an aperture resolution of 7×7 . Using NVIDIA’s CUDA implementation of BLAS, we achieve about 13 ms for the entire computation of one aperture pattern given an original image on an NVIDIA GeForce 8800 Ultra.

Using the computed aperture kernel, the original image then needs to be pre-compensated to counteract the optical defocus. While a compensation of dynamic content with a static kernel allows the inverse kernel to be pre-computed at different scales as described in section 5.2, these need to be computed at each frame for dynamically adapted aperture patterns. The additional computation times for a varying number of kernel scales is summarized in figure 10.

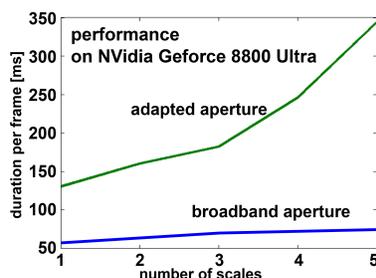


Fig. 10. While all required data for a static aperture pattern can be precomputed, the dynamically computed adaptive apertures add additional computational load to each frame, thus lowering performance for a larger number of kernel scales, i.e. a larger number of different distances to the focal plane.

Figure 11 shows several different images that are compensated using our adaptive apertures. Each of the examples shows the in-focus and defocused projection, as well as the compensated defocused case. The sub-images show the corresponding continuous (sec. 6.2) and binary (sec. 6.5) aperture patterns. Since our prototype is limited to binary patterns, only the latter are displayed on the integrated LCA. Although natural images roughly follow a $1/f$ frequency distribution, figure 11 shows that our adaptive apertures differ from a $1/f$ distribution depending on the image content and greatly enhanced the quality of a defocused projection.

7. EVALUATION AND COMPARISON TO PREVIOUS WORK

Due to the fact that our adaptive coded apertures are computed based on limitations of the human visual system we validate our theory using a visual difference predictor (VDP) [Mantiuk et al. 2005]. This operator can be applied to defocused and compensated images given in absolute luminance values under specific viewing parameters including adaptation luminance, distance to screen, and size of the canvas. The result is a map that contains the detection probabilities for each pixel compared to a reference image. We applied the VDP to our worst result (as explained in figure 13), the “house” scene in figure 11. Figure 12 visualizes the probability of difference detection for the defocused image, the compensated defocused image projected with a circular aperture, and a compensation with our static and dynamic apertures, compared to the focused original image. All input photos for the VDP were taken with the same camera settings in a calibrated environment. The probabilities are overlaid on the original image showing a high detection probability in red, moderate probabilities in green, and low ones in the original grayscale image value. The adaptive coded aperture performs best.

All previous work that attempted to compensate defocus with single display devices and circular apertures [Brown et al. 2006; Oyamada and Saito 2007; Zhang and Nayar 2006] achieves results that are similar to

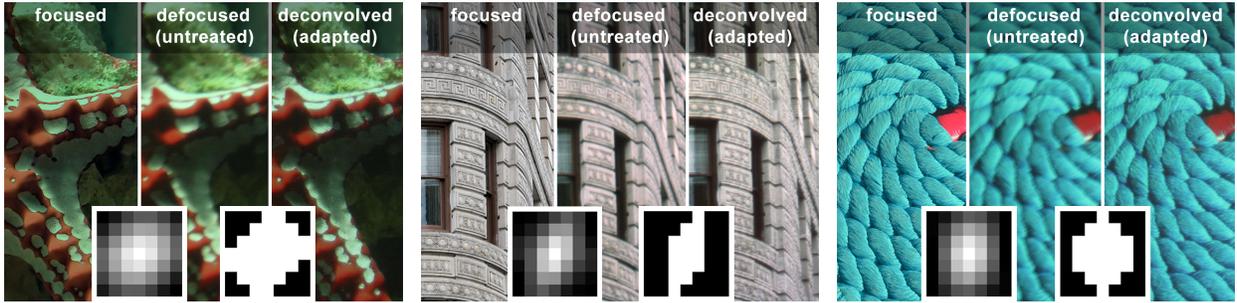


Fig. 11. Placing a transparent liquid crystal array at the aperture plane of a projector lens allows encoding the aperture’s mask pattern dynamically – depending on the perceivable frequencies of the displayed images. The inlays illustrate the computed intensity code and the applied binary masks. Each of the three image triplets are perception optimized and have been computed for viewing at a minimal distance of 50 cm when being displayed at a maximum diagonal of 21 cm on the screen. Possible artifacts could only be perceived when observing at closer distances or larger sizes.

the compensated case with circular apertures (figures 1 and 12). Although all of them slightly differ in their exact solution, the best achievable quality is limited by the physical aperture itself and is thus represented by a compensation with circular apertures in this paper.

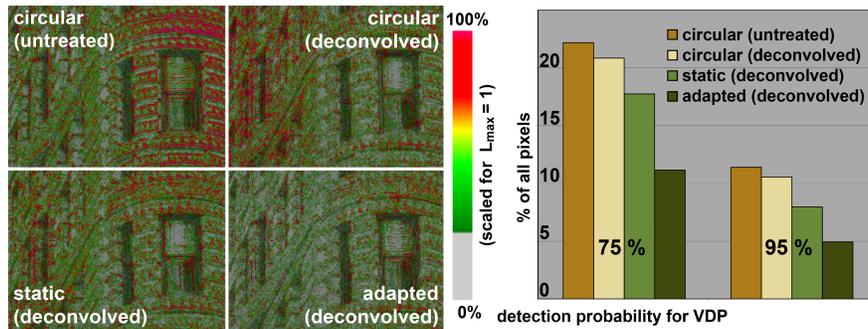


Fig. 12. Results of visual difference prediction when compared with focused projection: color coded detection probability (left) and bar charts for 75% and 95% difference detection probabilities (right) – all for house example in figure 11.

For evaluation we also wish to quantitatively compare the depth-of-field that we can achieve to previous work and conventional projection, as well as the loss in light transmission. 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 objective’s f -stop number. A low f -stop is equivalent to a large light throughput with a small depth-of-field and vice versa. A sufficiently wide depth-of-field is necessary for projections onto curved surfaces, such as domes or cylindrical screens. However, since bright images are desired for almost all applications, diaphragms that reduce the effective aperture diameter in favor to a wider depth-of-field are normally not incorporated into projectors. Exceptions are devices with automatic iris control (auto-iris) that can dynamically adjust the diameter of a circular diaphragms and, thus, control the DOF.

Following Oyamada and Saito [2007], 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

	defocused (untreated)	circular (deconvolved)	broadband (deconvolved)
$\hat{f}/\#$	3.1	3.7	4.3
$\tilde{f}/\#$	3.1	3.1	3.9

	adapted (deconvolved)			
	house	starfish	rope	lena
$\hat{f}/\#$	4.4	5.4	5.6	7.7
$\tilde{f}/\#$	4.0	3.5	3.7	3.7

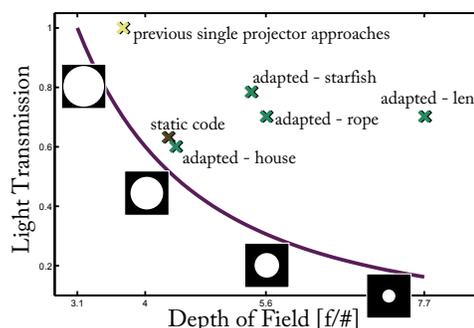


Fig. 13. The table on the left shows a comparison of depth-of-field enhancement as well as light throughput of compensated projections with our coded and circular apertures. The values are given in relative f-numbers comparing DOF ($\hat{f}/\#$) and transmission ($\tilde{f}/\#$). Higher values are better for DOF enhancement, lower values better for light transmission. Both, DOF and light transmission increase from circular to broadband and even more to adapted apertures. The plot on the right illustrates the tradeoff between DOF and transmission for the examples in the table and for an auto-iris system (curve).

aperture (and constant focal length) that would lead to the same depth-of-field ($\hat{f}/\#$). For the “lena” image in figure 1, for instance, an $\hat{f}/7.7$ aperture stop is needed to achieve the same DOF with a standard projector, which corresponds to a tremendous amount of light loss. In terms of light throughput, the gain factor is theoretically $\times 6.1$. Due to the limited contrast of our LCA, however, it is approximately $\times 2.5$ in practice. As seen in figure 13, previous approaches using deconvolution with circular apertures can slightly enhance the DOF without losing any light, whereas our static broadband mask and especially the adaptive codes achieve a much better gain in DOF. The enhanced DOF comes at the cost of reduced light transmission, which is the case for all the examples in figure 13 (left). We evaluate the light throughput also in f-numbers ($\tilde{f}/\#$) with respect to a comparable standard projector equipped with an auto-iris. The results show that our patterns can always achieve a larger DOF without losing as much light as an auto-iris projector for which $\hat{f}/\# = \tilde{f}/\#$. Our adaptive codes also always achieve larger DOFs than the static coded aperture, while maintaining a higher light throughput.

8. OTHER APPLICATIONS FOR CODED APERTURES

So far, we have seen how static and dynamic coded apertures can improve the depth-of-field of a projection system, while maintaining a high light transmission. Our system is, however, more powerful and flexible than that. In the following, we outline two different applications of the proposed projection systems: smooth focused projections via de-pixelation and improved temporal image contrast.

Pixelation is an artifact created by all digital projectors. Due to spacing between the display pixels and their spatial discretization, focused projections often appear jaggy as shown in figure 14. This effect becomes especially irritating for high resolution photographs of projected imagery, as modern digital cameras have an ever increasing spatial resolution. Following Zhang and Nayar [2006], we can introduce a slight optical defocus that lets the pixel structures optically vanish. Using our defocus compensation, we can invert the effect and thus display sharp images without visible pixel structures. Results for a static aperture de-pixelation (scene from figure 8) are shown in figure 14 in the upper row, while the lower row presents a closeup of the adaptive “rope” scenario of figure 11.

Another application of our system is improved temporal contrast. Modern auto-iris projection systems automatically adjust the iris opening to dynamically control the amount of projected light. For video

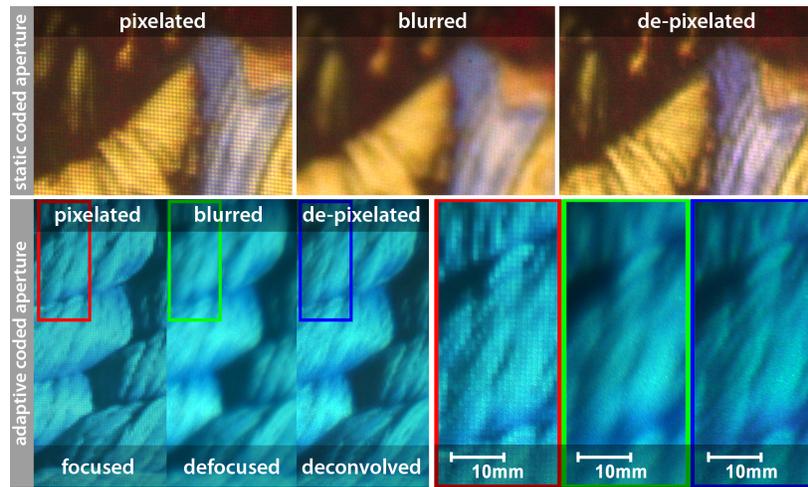


Fig. 14. In some situations, the display pixel grid can become visible and disturbing for focused projections. A slight optical defocus diminishes the visibility of the grid structure, but also blurs image features. Using our coded apertures and corresponding defocus compensation algorithms, we can achieve smooth and sharp projections as shown in these examples. The upper row shows magnifications for a de-pixelation performed with our static broad-band mask, while the lower row shows results from our adaptive aperture.

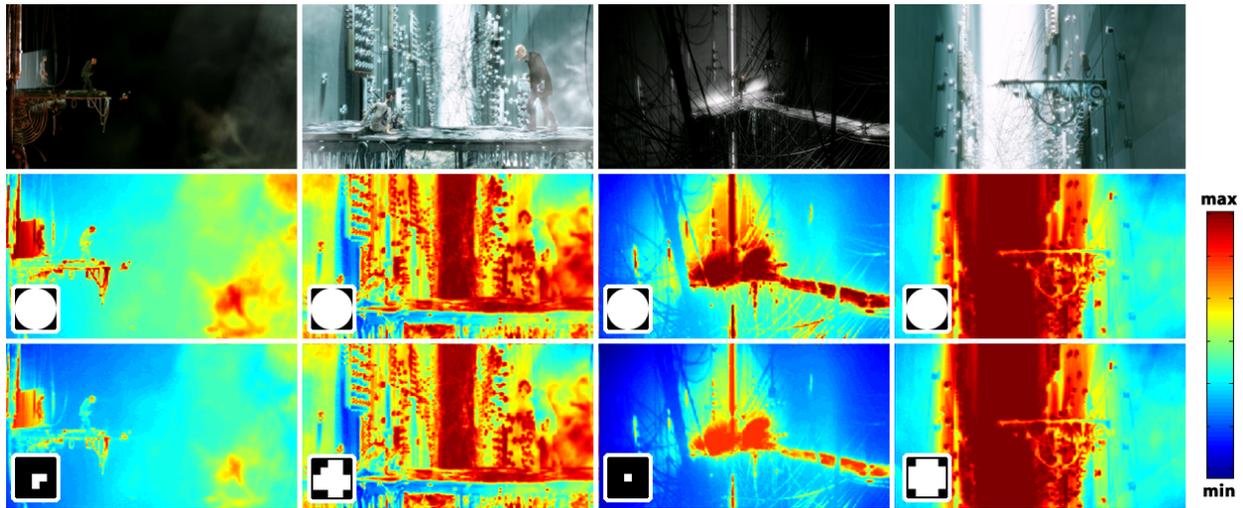


Fig. 15. The temporal contrast of displayed video sequences can be increased using adaptive coded projector apertures. For this purpose, the aperture pattern is adjusted in accordance with the presented image brightness. For very bright frames (columns 2 and 4) this yields an open aperture, while dark images result in small aperture openings (columns 1 and 3), thus effectively reducing the projector's blacklevel. Four frames of a video sequence are shown in the upper row, while rows 2 and 3 are logarithmically color-coded luminance values with unmodified and scaled coded apertures, respectively. In contrast to standard auto-iris lenses, luminance-scaled adaptive coded apertures additionally increase the depth-of-field.

sequences with frames that differ significantly in image brightness, this leads to an increase in inter-frame or temporal contrast. Specifically, the advantage for auto-iris displays is the option to either transmit all possible

light for very bright images, while reducing the aperture size for dark frames. The projector blacklevel can be reduced and the temporal contrast therefore increased. Using our adaptive aperture projector, we can adjust the attenuation pattern to control the light throughput, just like auto-iris systems. This is demonstrated in figure 15, where the aperture opening is scaled according to the average image brightness for several frames of a video sequence (after being computed as explained in section 6). The projected frames are color-coded for better visualization and show how the blacklevel is decreased for dark images.

9. DISCUSSION

We have presented algorithms and prototypes to compensate optical defocus of a single video projector by means of inverse light filtering and coded apertures. We have shown how integrated aperture patterns can increase the depth-of-field of ordinary video projectors. Adaptive coded apertures optimize depth-of-field versus light throughput based on the limitations of the human visual system, and can potentially lead to a new generation of auto-iris projector lenses. They outperform static broadband masks and circular aperture stops for digital projector defocus compensation and lead to a larger depth-of-field, higher quality projector de-pixelation, and increased temporal contrast of projected video sequences. Static coded apertures, however, are easy to manufacture and to integrate into existing projector designs, and are therefore more economical. In particular, projector de-pixelation, where the discretized pixel structure is optically diminished through defocus while image details are recovered through inverse filtering (see figure 14), will not only improve the image quality of close-view projection-based displays (such as rear-projected TV sets) but also of devices that utilize projection-based illumination techniques. High-dynamic-range displays that apply spatially modulated projected backlight, such as [Seetzen et al. 2004], suffer from visible moiré patterns if the projection is focused on the screen plane. To avoid this, HDR displays usually apply LED-based backlights. Using our approach, the low contrast frequency imposed by low-frequency backlights could be significantly improved. The hardware and software solutions we have presented in this paper could form an integral part of next generation projection devices.

The main limitations of our adaptive coded apertures are currently imposed by the employed LCA. Its low light transmittance (only 30% when completely transparent) results in a relatively high loss of light. Therefore, we trade light throughput for depth-of-field. Furthermore, the maximal contrast of super-twisted-nematic LCAs (as ours) is currently 7:1. As spatial light modulators (SLMs), such as a high contrast continuously valued LCAs with higher transmittance, reflective digital micromirror devices (DMD), and liquid crystal on silicon (LCoS) chips, become more widely available we expect better results with these displays. DMDs also do not suffer from the light loss and low contrast as LCDs do. We have shown that our approach is feasible and plan to experiment with alternative SLMs in the future. We also believe that high contrast and brightness at low power consumption and heat development will become feasible with light engines that apply upcoming LED technology.

Another limitation of our current adaptive aperture prototype is the moderate performance for a large number of compensated defocus scales. This will improve with next generation graphics hardware, or with customized integrated image processors. With a higher hardware performance, larger resolution aperture patterns can be computed. This would lead to less constraints for the discrete scaling operations that are required for inverse filtering and temporal aperture adaptations for a flicker-free increased temporal contrast. Currently, we are constrained to computations with an aperture resolution of 7x7 pixels.

Our adaptive coded aperture prototype is limited to binary patterns, therefore, we have to binarize our adaptive masks. Although this rounding and other approximations make the eventually displayed aperture codes not quite optimal in practice, we believe that the quality of the presented results and the generality of our approaches, which also apply for alternative prototypes or commercial displays, demonstrate the feasibility of our algorithms and hardware setups.

In the future, we would like to experiment with more sophisticated models for frequency perception of the human visual system. Furthermore, exploring perception-optimized numerical minimization approaches is an interesting avenue for future research. Although a standard least-squared error solution to the defocus compensation problem achieves very good results, it is most likely not the optimal solution for a human observer even if additional constraints such as non-negativity would be included.

REFERENCES

- AHUMADA, A. J. AND WATSON, A. B. 1985. Equivalent-Noise Model for Contrast Detection and Discrimination. *J. Opt. Soc. Am. A*, 2, 1133–1139.
- BANDO, Y., CHEN, B.-Y., AND NISHITA, T. 2008. Extracting Depth and Matte using a Color-Filtered Aperture. *ACM Trans. Graph. (Siggraph Asia)* 27, 5, 1–9.
- BIMBER, O. AND EMMERLING, A. 2006. Multifocal Projection: A Multiprojector Technique for Increasing Focal Depth. *IEEE TVCG* 12, 4, 658–667.
- BROWN, M. S., SONG, P., AND CHAM, T.-J. 2006. Image Pre-Conditioning for Out-of-Focus Projector Blur. In *Proc. IEEE CVPR*. Vol. II. 1956–1963.
- DALY, S. 1993. The Visible Differences Predictor: An Algorithm for the Assessment of Image Fidelity. In *Digital Image and Human Vision*, A. Watson, Ed. Cambridge, MA: MIT Press, 179–206.
- DOWSKI, JR, E. AND CATHEY, W. 1995. Extended Depth of Field through Wavefront Coding. *Applied Optics* 34, 11, 1859–1866.
- DOWSKI, E. R. AND CATHEY, W. T. 1998. Apparatus and Methods for Extending Depth of Field in Image Projection Systems . US patent 6069738.
- DURAND, F. AND DORSEY, J. 2000. Interactive Tone Mapping. In *Proc. Eurographics Workshop on Rendering*. 219–230.
- GROSSE, M. AND BIMBER, O. 2008. Coded Aperture Projection. In *IPT/EDT*. 1–4.
- GRUNDHOFER, A. AND BIMBER, O. 2008. Real-Time Adaptive Radiometric Compensation. *IEEE TVCG* 14, 1, 97–108.
- ITU. 1993. Iso/iec 10918-1: 1993(e) ccit recommendation t.81.
- KRAWCZYK, G., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Perceptual Effects in Real-Time Tone Mapping. In *Spring Conference on Computer Graphics 2005*. ACM, 195–202.
- LEVIN, A., FERGUS, R., DURAND, F., AND FREEMAN, W. T. 2007. Image and Depth from a Conventional Camera with a Coded Aperture. *ACM Trans. Graph. (Siggraph)* 26, 3, 70.
- LEVIN, A., HASINOFF, S. W., GREEN, P., DURAND, F., AND FREEMAN, W. T. 2009. 4D Frequency Analysis of Computational Cameras for Depth of Field Extension. *ACM Trans. Graph. (Siggraph)* 28, 3, 1–14.
- LIANG, C.-K., LIN, T.-H., WONG, B.-Y., LIU, C., AND CHEN, H. 2008. Programmable Aperture Photography: Multiplexed Light Field Acquisition. *ACM Trans. Graph. (Siggraph)* 27, 3, 55:1–55:10.
- MANTIUK, R., DALY, S. J., MYSZKOWSKI, K., AND SEIDEL, H.-P. 2005. Predicting Visible Differences in High Dynamic Range Images - Model and its Calibration. In *Proc. IS&T/SPIE's Annual Symposium on Electronic Imaging*. Vol. 5666. 204–214.
- MORENO-NOGUER, F., BELHUMEUR, P., AND NAYAR, S. 2007. Active Refocusing of Images and Videos. *ACM Trans. Graph.* 26, 3, 67.
- NG, R. 2005. Fourier Slice Photography. *ACM Trans. Graph. (Siggraph)* 24, 3, 735–744.
- OYAMADA, Y. AND SAITO, H. 2007. Focal Pre-Correction of Projected Image for Deblurring Screen Image. In *Proc. IEEE ProCams*.
- RAMASUBRAMANIAN, M., PATTANAIK, S. N., AND GREENBERG, D. P. 1999. A Perceptually Based Physical Error Metric for Realistic Image Synthesis. *ACM Trans. Graph. (Siggraph)* 25, 3, 73–82.
- RASKAR, R., AGRAWAL, A., AND TUMBLIN, J. 2006. Coded Exposure Photography: Motion Deblurring using Fluttered Shutter. *ACM Trans. Graph. (Siggraph)* 25, 3, 795–804.
- SEETZEN, H., HEIDRICH, W., STUERZLINGER, W., WARD, G., WHITEHEAD, L., TRENTACOSTE, M., GHOSH, A., AND VOROZCOVS, A. 2004. High Dynamic Range Display Systems. *ACM Trans. Graph. (Siggraph)* 23, 3, 760–768.
- VEERARAGHAVAN, A., RASKAR, R., AGRAWAL, A., MOHAN, A., AND TUMBLIN, J. 2007. Dappled Photography: Mask Enhanced Cameras for Heterodyned Light Fields and Coded Aperture Refocusing. *ACM Trans. Graph. (Siggraph)* 26, 3, 69.
- WETZSTEIN, G. AND BIMBER, O. 2007. Radiometric Compensation through Inverse Light Transport. In *Proc. Pacific Graphics*. 391–399.
- ZHANG, L. AND NAYAR, S. K. 2006. Projection Defocus Analysis for Scene Capture and Image Display. *ACM Trans. Graph. (Siggraph)* 25, 3, 907–915.

- ZHOU, C. AND NAYAR, S. K. 2009. What are Good Apertures for Defocus Deblurring? In *IEEE International Conference on Computational Photography*. 1–8.
- ZOMET, A. AND NAYAR, S. 2006. Lensless Imaging with a Controllable Aperture. In *IEEE CVPR*. 339–346.