Highlighted Depth-of-Field Photography: Shining Light on Focus

JAEWON KIM
MIT Media Lab and Korea Institute of Science and Technology(KIST)
ROARKE HORSTMEYER
MIT Media Lab
IG-JAE KIM
MIT Media Lab and Korea Institute of Science and Technology(KIST)
and
RAMESH RASKAR
MIT Media Lab

We present a photographic method to enhance intensity differences between objects at varying distances from the focal plane. By combining a unique capture procedure with simple image processing techniques, the detected brightness of an object is decreased proportional to its degree of defocus. A camera-projector system casts distinct grid patterns onto a scene to generate a spatial distribution of point reflections. These point reflections relay a relative measure of defocus that is utilized in postprocessing to generate a highlighted DOF photograph. Trade-offs between three different projector-processing pairs are analyzed, and a model is developed to help describe a new intensity-dependent depth of field that is controlled by the pattern of illumination. Results are presented for a primary single snapshot design as well as a scanning method and a comparison method. As an application, automatic matting results are presented.

Categories and Subject Descriptors: I.3.3 [**Computer Graphics**]: Picture/ Image Generation—*Viewing algorithms*; I.4.1 [**Image Processing and Computer Vision**]: Digitization and Image Capture

General Terms: Algorithms, Design

Additional Key Words and Phrases: Computational photography, HDOF photo, depth of field, active illumination, matting, image processing

This research was conducted with the MIT Media Lab's Camera Culture group. R. Raskar was supported by an Alfred P. Sloan Research Fellowship.
Authors' addresses: J. Kim (corresponding author), R. Horstmeyer, I.-J. Kim, and R. Raskar, MIT Media Lab, 77 Mass. Ave., E14/E15, Cambridge, MA 02139-4307; email: jaewonk@media.mit.edu.

DOI 10.1145/1966394.1966403
http://doi.acm.org/10.1145/1966394.1966403

ACM Reference Format:
Kim, J., Horstmeyer, R., Kim, I.-J., and Raskar, R. 2011. Highlighted depth-of-field photography: Shining light on focus. ACM Trans. Graph. 30, 3, Article 24 (May 2011), 9 pages. DOI = 10.1145/1966394.1966403 http://doi.acm.org/10.1145/1966394.1966403

1. INTRODUCTION

A common technique in photography is to use the limited depth of field of a lens to emphasize and frame focused objects while deemphasizing the rest of a scene. Photographers often use expensive, large aperture lenses to achieve this blur effect in macrophotography and portraits. We present a camera setup that can decrease the brightness of out-of-focus objects, providing an additional tool for photographers to achieve their composition goals. The output of this camera is called a Highlighted Depth Of Field (HDOF) photo (Figure 1). Specifically, the design uses a projector to display point patterns on a particular scene, and resamples or combines images to achieve the desired intensity shift.

1.1 Contributions

We present an analysis of camera-projector setups that change the apparent brightness of objects based on their distance from the camera's focal plane. Creating an intensity gradient along the z-dimension could be useful for object segmentation, contrast enhancement, or simply for creative effects. Included in this analysis are the following:

—a geometric and physical optics model of defocus for a projected grid pattern, which establishes a method to decrease the brightness of out-of-focus objects,
—three unique projection-processing methods to create HDOF photographs, including a single-shot method, a two-shot method, and a multishot method, each presenting a unique trade-off between required number of images and final image resolution,
—Example applications, including high-frequency feature segmentation and depth-range-selectable matting techniques.

1.2 Related Work

Following is a brief overview of relevant imaging systems that use illumination to assist in the segmentation of depth information, which is summarized in Figure 2.

(a) Conventional Photo (b) Successive dimming of defocused areas highlights objects within the camera's depth of field.

Fig. 1. Can we create photos where out-of-focus objects are not only blurred but also dimmed? Modern projector-as-flash accessories may provide an opportunity for new focus-related artistic effects. Top photos compare a conventional photo to a highlighted DOF photo when a woman is in-focus and the doll is slightly out-of-focus. The doll is successively darkened from left to right while the brightness of the woman is maintained. The bottom photos show a similar comparison with the camera focused in an opposite configuration.

Item \ Method	Range Camera	Stereo Camera	Light field Camera	Highlighted DOF
Depth Cue	Time of Flight	Parallax	Refocusing Pixel Var.	Reduced Intensity
Cost	Very High	Cheap	Medium	Cheap
Dealing with Specularity	Poor	Good	Poor	Poor
Outdoors	Poor	Very Good	Very Good	Poor
Motion	Good	Good	Good	Medium
Areas with no Texture	Good	Poor	Poor	Very Good
Final Resolution	Good	Good	Poor	Very Good
Depth dependent extinction of brightness	Good	Medium	Good	Very Good

Fig. 2. A comparison of different possible setups for obtaining depth content. A highlighted DOF photo could be generated by first estimating depth or creating focal stacks via a light field, and then dimming appropriate layers. Intead, our method uses projected patterns to maintain a high spatial resolution, and can even work in flat areas with no texture.

Scene Geometry and Structured Lighting. Investigations into using projected light to infer about a scene's geometry began with Will and Pennington [1971]. Since, projected patterns have found a wide array of applications in the estimation of 3D object shape [Mouaddabi et al. 1997; Schechner et al. 2000; Salvi et al. 2004; Davis et al. 2005]. Also, projected line and grid patterns have been used to determine surface orientation [Wang et al. 1987; Shrikhande and Stockman 1989; Maas 1992], to separate direct and global light components [Nayar et al. 2006; Kim et al. 2010], and to assist in robot mobility [LeMoigne and Waxman 1988]. While we also use a projected array of points to highlight the plane of focus, the final goal is not to estimate complex scene characteristics. Instead, we characterize depth information by optical modulation of any portion of the scene that is not in sharp focus on the sensor.

Confocal Microscopy. The fundamental principle of a confocal microscope, which uses a pinhole to distinguish light from a particular depth, is very similar to the basis of our system. The confocal design has also been modified to use a more efficient array of microlenses [Tiziani and Uhde 1994] or pinholes [Eisner et al. 1998], which we experimented with over our sensor. In general, smaller pinholes lead to a narrow depth of field and higher axial resolution in confocal designs. Following the same logic, we project small spots to better create a narrow depth of field around the focal plane. There has been recent work to incorporate light field imaging with confocal microscopy, as well as light field capture with a 4D illumination source from Levoy et al. [2004, 2006, 2009]. These setups can capture multiple planes of focus in a single image and share the common goal of highlighting a certain depth plane. Conceptually similar examples using time-varying illumination are also found in confocal literature. Mitic et al. [2003] implement a time-varying grid pattern in a wide-field microscope to obtain depth discrimination by comparing multiple images. Wilson et al. [1996] present a unique aperture correlation technique to increase optical efficiency by illuminating and imaging through a scanning random mask. Finally, Heintzmann et al. [2001] incorporate a micromirror array to image and illuminate along two optical paths, which increases conventional SNR twofold. These procedures overlap with the basic concept of our multishot methods, but require uniquely different illumination patterns and processing algorithms.

Depth from Defocus. A number of papers and software packages, like Depth of Field Generator Pro, offer methods of detecting or creating additional defocus through computational means by first creating a depth map [Lai et al. 1992; Bae and Durand 2007]. Other computational-based depth estimation methods compare images from different viewpoints, as in stereovision, or compare images from a single camera with a variable aperture [Watanabe and Nayar 1998; Hasinoff and Kutulakos 2006]. Attempting to acquire scene depth with a projector system has been demonstrated before using multiple images by Zhang and Nayar [2006]. Furthermore, full resolution refocusing from a single image is achieved with a projected

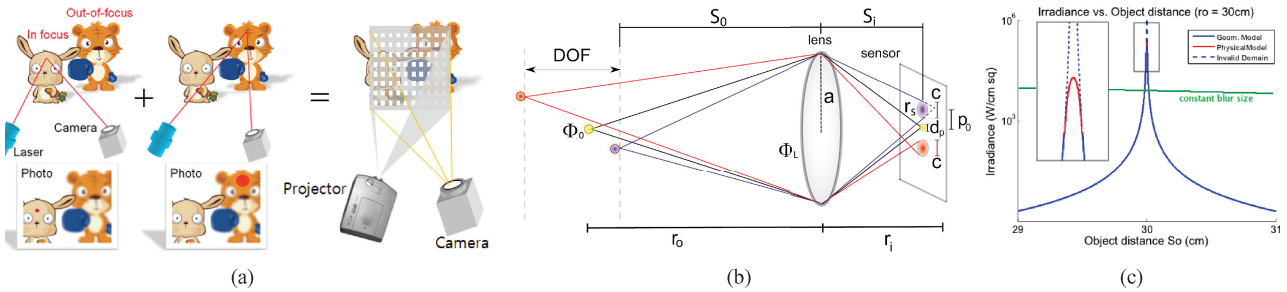


Fig. 3. (a) A dot from a laser pointer viewed through a camera will lose localized intensity as it becomes defocused. We use a large DOF projector to create a dense array of dots. (b) Imaging three point sources at different focal depths, with setup variables labeled. (c) A plot of expected irradiance vs. object distance for an $f = 5$cm camera focused at 30cm, where the object in question is an idealized 1W point source. The line in green also shows the intensity of a hypothetical imaged spot that does not change in size (i.e., it remains at $10\mu m$), to indicate that defocus plays the dominant role in specifying image intensity.

dot pattern in Moreno-Noguer et al. [2007]. This is essentially a depth estimation method that works only for scenes with low depth complexity (i.e., scenes with little texture that can be easily segmented), as the depth is estimated at sparse locations with a delicate calibration process. Such an approach cannot handle segmentation at complex geometric boundaries. One of our methods, the single-shot method, generates real-valued depth cues for each pixel at 1/9th the full resolution without any calibration. The other two proposed methods provide full resolution depth cues, which can be applied to create nonbinary mattes, for example, using two or more images.

2. ANALYSIS OF INTENSITY VARIATION WITH DEFOCUS

This section begins with a simplified derivation of a defocus-intensity relationship, which helps explain the basic concept of our design. We then extend this relationship to determine optimal projection patterns and processing methods to create an HDOF image in Section 3.

2.1 Point Source Model

A generalized model that analyzes the behavior of our camera-projector system with respect to the location of a point source will help explain its utility. Figure 3(b) depicts the geometric effect of defocus for a simplified camera. Our analysis will begin by considering the imaging of spatially separated point sources of light. An ideal point source with an initial radiant flux Φ_0 emits light in all directions, with the total radiant flux reaching a lens of radius a at distance s_o given by $\Phi_{lens} = \Phi_0(a^2/4s_o^2)$, from the intensity law of geometrical optics. In practice we use a projector to create this point source, causing it to slightly increase in size and decrease in brightness with depth, which we will ignore for simplicity. As we will see, the dominant change in the point source's detected intensity will be a result of defocus. The size of the imaged point source's area will depend on how in-focus it is. For a circular lens, under rough geometric optics assumptions of large defocus, the majority of the point's light will lie within a circle of radius r_s on the sensor. This radius can be expressed in terms of known camera dimensions, through the use of similar triangles, as $r_s = a\,|1 - r_i/s_i|$, where r_i is the distance to the sensor plane, s_i is the distance at which there is a sharp image (i.e., plane conjugate for the object distance s_o), and the absolute value takes into account both positive and negative defocus. The quantities r_i and s_i are related to the object distances r_o and s_o by the thin lens law, $r_i = r_o f/(r_o - f)$, where f is the focal length. With the radius r_s, the irradiance at a given point in the blur spot can be estimated as a function of object distance

$$I_s(s_o) = \frac{\Phi_{lens}}{\pi r_s^2} = \frac{\Phi_0}{4\pi s_o^2}\left|1 - \frac{r_o(s_o - f)}{s_o(r_o - f)}\right|^{-2}, \quad (1)$$

assuming a constant irradiance over the blur area. Eq. (1) is plotted in Figure 3(c) for typical camera parameters in blue, and it is clear that this geometric model approaches an asymptote for an in-focus point source. A physical optics approximation for intensity under small amounts of defocus can be used to correct the geometrical model. A PSF's peak intensity is commonly determined with $I_{max} = \Phi_{lens}(\frac{\pi a^2}{\lambda^2 f^2})$, and drops as [Born and Wolf 1970]

$$I_s \approx I_{max}(1 - (2\pi/\lambda)^2 \Delta W^2), \quad (2)$$

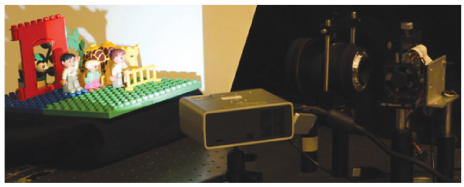
where ΔW^2 is the mean square wavefront deformation at the pupil plane. Combining these expressions yields a physical optics prediction of irradiance as a function of object distance

$$I_s(s_o) = \Phi_0\left(\frac{\pi a^4}{4{s_o}^2\lambda^2 f^2}\right)\left[1 - \frac{\pi^2}{12\lambda^2 {r_i}^4}\left(\frac{s_o f}{s_o - f} - r_i\right)^2\right], \quad (3)$$

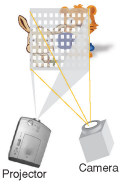
which is also plotted in Figure 3(c) in red. To be clear, Eq. (3) includes the drop in detected intensity due to the point's distance. For a single point source that is within a few meters, the detected intensity at a given pixel drops much quicker as a function of defocus than as a function of object distance. This is clear from Figure 3(c), where intensity sharply but asymmetrically drops from the in-focus plane, remaining marginally higher for close objects. This drop is not typically noticeable for images of a continuous scene, as blurred rays simply overlap and integrate to a higher value. By projecting a spatial grid of bright point sources, however, we aim to artificially create this effect. Again, due to the finite size of projector pixels, our grid will not contain perfect point sources. Specific parameters of this grid will now be determined and connected to depth of field.

2.2 Intensity-Dependent Depth of Field

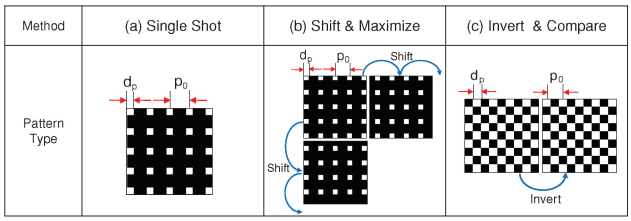
Given a characterization of depth from the intensity of a single defocused point source, we can now describe how a distributed illumination pattern can create an arbitrarily narrow artificial depth of field. To begin, a conventional camera at a particular lens setting has a depth of field defined by an acceptable circle of confusion c. Referring to Figure 3(b), we can use the focal length relationship



(a) Actual setup



(b) Diagram



(c) Projection patterns

Fig. 4. A summary of the setup and projection patterns for our three HDOF photography designs.

and similar triangles to find that

$$DOF = afr_i\left(\frac{1}{ar_i - f(a+c)} - \frac{1}{ar_i - f(a-c)}\right), \qquad (4)$$

where DOF represents depth of field of the camera in object space, and c is typically chosen to be a blur spot value less than the size of a sensor pixel. As noted in the previous section, as long as a point source of light is isolated, its detected intensity decreases with defocus. Once light from blurred spots begin to overlap, intensity values will integrate towards increasing values. Following this concept, we can define the acceptable blur spot size c for Eq. (4) as the distance between the center of two projected points in image space, x_1 and x_2, as $c = |x_1 - x_2| = p_o$. Here, p_o is the constant pitch between projected points as seen on the camera sensor. Plugging this value for c into Eq. (4), it becomes clear that a narrower intensity gradient for a fixed camera setup can be generated with an increase in pattern pitch. We call this an intensity-dependent DOF. We can also use the pattern pitch to find the maximum drop in intensity with defocus for a point source in the grid

$$\delta I = I_{max} - I_{min} = I_{max}(1 - (d_p/p_o)^2), \qquad (5)$$

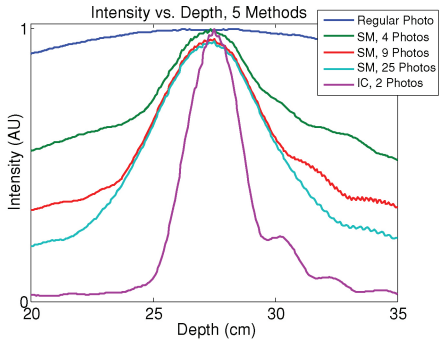
where I_{max} is given from Eq. (3). Here, d_p is the finite width of the in-focus projected spot, which is ideally the width of one camera pixel. In the following section, we will present methods using multiple images that allow for more control over the intensity gradient.

3. PROJECTOR AND PROCESSING SETUPS

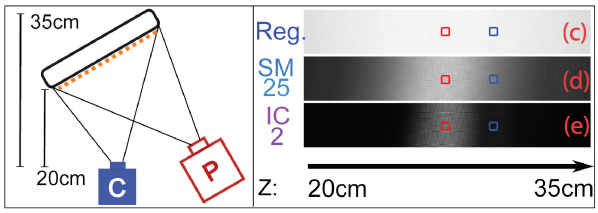
This section outlines the specific attributes of three projection patterns and image processing techniques based on the concepts from Section 2. A summary of all methods is in Figure 4, with example results in Figure 6 and Figure 7.

3.1 Single-Shot Method

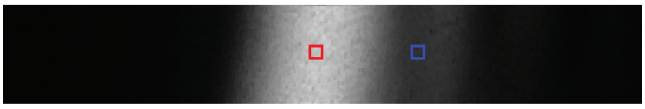
The primary implementation of our imaging technique records a spatial distribution of point patterns on a single photograph. A simple physical setup to achieve this uses a projecting element and camera. The element could be a laser-grating pair, a uniquely designed flash, or an off-the-shelf projector, as long as it has a sufficiently wide depth of field. We use a digital projector placed close



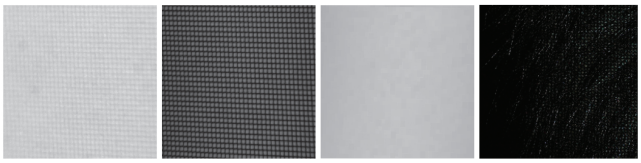
(a) Experimental traces of different projector-processing combinations, which create a variety of sharp intensity gradients at the plane of focus.



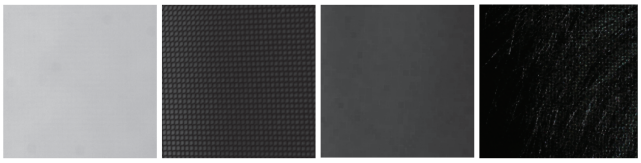
(b) The experimental imaging setup for the intensity plots in (a), as well as cropped raw photographs for (c) a conventional image under full illumination, (d) a Shift and Maximize image, and (e) for an Invert and Compare image.



(f) Result of a morphology operation on (e)

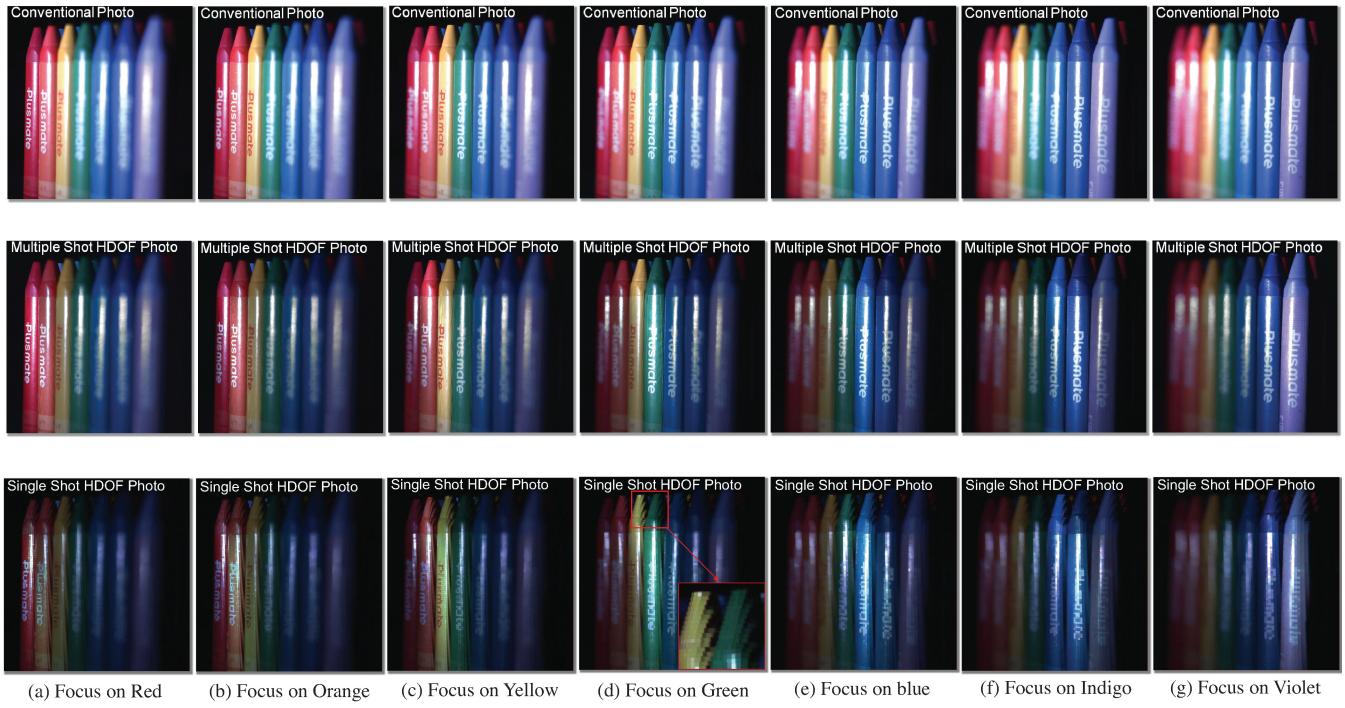


(g) Magnification of red-boxed area (the in-focus region) for (c) on left, (e) in middle, and the morphology operation (f) on right. An example raw variance map is shown at the far right.



(h) Magnification of blue-boxed area (slightly defocus region) for (c) on left, (e) in middle, and the result of the morphology operation on (f) on right. The variance map created after the morphology operation, on far right, shows little change from (g) right.

Fig. 5. An experimental demonstration of the different HDOF techniques performed by imaging a tilted Lambertian object. (a)–(f) Different techniques can create different intensity gradients. (g)–(h) Artifacts can be corrected using a simple closing operation applied globally to the image with minimal loss to depth sensitivity.



(a) Focus on Red (b) Focus on Orange (c) Focus on Yellow (d) Focus on Green (e) Focus on blue (f) Focus on Indigo (g) Focus on Violet

Fig. 6. The top row displays 7 images of a tilted crayon box, each focused at a slightly different depth plane using a large aperture setting (f/2.8) for a close-up (15cm) scene. (Middle) Results from the multishot method. (Bottom) Results from the single-shot method, which exhibit pixelation and artifacts due to imperfect sampling.

(a) Conventional (b) Multi-Shot (c) Two-Shot

(d) Conventional (e) Multi-Shot (f) Two-Shot

Fig. 7. Comparison between original photos, multishot HDOF photos, and two-shot HDOF photos. In (a,b,c) the focus is on the front objects, while in (d,e,f) the focus is on the middle objects. Note the brightness change in the background male doll in (a,b,c), while there is no brightness change in the foreground female doll. From left to right, the average pixel value across the background doll's face is 84.4, 44.7, and 49.2, while the average value across the foreground doll's face is 110.5, 109.5, and 110.0. In (d,e,f) the average value across the foreground doll's face is 111.5, 64.6 and 53.5 while the average value in the background is 83.0, 82.8 and 82.5.

to a camera for our primary design. We have already introduced the basics of how the peak intensity of a projected point delivers a relative measure of defocus. An output image with dimmed out-of-focus regions is created by sampling the original image at the center of each projected point. These centers are located by searching for local maxima in small areas. This assumes that the surface texture does not contain high spatial frequencies. Total spatial resolution, maximum intensity drop, and the intensity-dependent-*DOF* width will all be variably set by the projection pattern duty cycle. While we have currently settled on an optimal pitch to balance these three parameters (1/9 duty cycle, i.e., every 3rd pixel in each row and column), we expect some future flexibility in pattern selection given the trend of ever-increasing projector and sensor resolution.

3.2 Multishot Method: Shift and Maximize

Instead of undersampling a single image to create an HDOF photo, a full resolution output can be generated by processing multiple images. It is clear that projecting a binary grid pattern will only illuminate a certain percentage of a scene. The multishot method simply applies the previously discussed concepts to several images of a shifted grid pattern, allowing the entire scene to be illuminated over time. Shifting can be achieved either digitally, as in our designs, or with mechanical motion. The number of shifted projection images, n^2, required for a fully illuminated scene is given by the inverse of the projector duty cycle: $n = p_o/d_p = c/d_p$. Here, the number of required images is also given in terms of c to highlight its relationship to the intensity-dependent *DOF* in Eq. (4). As the projector duty cycle is decreased and more images are taken, a narrower *DOF* and larger intensity gradient can be created. This concept is demonstrated with a simple experiment in Section 4.1. The most direct way to compute an HDOF image is by selecting the maximum intensity value for each pixel over the range of captured images. Representing the three-dimensional dataset of images as D_{ijk} with k denoting image index, this is expressed as

$$F_{ij} = max_k(D_{ijk}). \tag{6}$$

The multishot method proved to be the most flexible gradient-generating technique, but is best suited for static scenes.

To summarize the procedure for the multishot method:

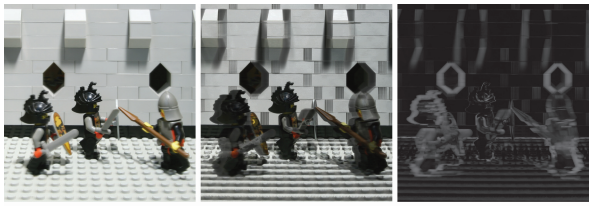
(1) project a grid pattern with duty cycle $(d_p/p_o)^2$ and take an image;
(2) shift the pattern a distance d_p in image space;
(3) repeat steps 1 and 2 n^2 times;
(4) from within all images take the max value for each pixel.

3.3 Two-Shot Method: Invert and Compare

Comparing two images of a binary shifted checkerboard projection pattern creates a robust statistical approximation of defocus. Combining these two images generates an image that appears to have been taken under constant white-light projector illumination. So, processing only two images generates a full resolution image exhibiting a high intensity gradient at the focal plane. Specifically, in-focus areas will contain a resolved checkerboard pattern that inverts in a second image (Figure 4(c)). The maximum intensity drop will only be a factor of 4 from Eq. (5), since $d_p = p_0/2$. For objects further away from the focal plane, the checkerboard pattern blurs to a constant intensity value that remains unchanged with pattern inversion. Taking the normalized absolute difference between the two images and multiplying it by their sum yields

$$F_{ij} = \left[\sum_{RGB} C\frac{|D_{ij1} - D_{ij2}|}{D_{ij1} + D_{ij2}}\right]\left(D_{ij1}^{RGB} + D_{ij2}^{RGB}\right), \tag{7}$$

where C is a constant, the defocus variance map is summed over color channels, and the image sum is not. The resulting intensity-dependent *DOF* using this comparison algorithm is sharper than the multishot method. Specifically, since each illuminated point can only blur to roughly twice its original size before overlapping with adjacent points and approaching a constant value, the variance map quickly approaches zero for defocused regions. Thus, this technique is best suited for an application where one wishes to highlight a narrow plane in full resolution, as in a fixed macrophotography setting.



(a) refocusing (b) LF HDOF photo (c) Variance map

Fig. 8. Light Field (LF) camera [Ng et al. 2004] or camera-array-based methods [Joshi et al. 2006] provide focus stacks but do not provide depth cues in flat regions with no texture. (a) A refocused light field image, focusing on the middle toy. (c) A variance map with high variance in regions of defocus. Errors in the irregular bright patterns (i.e., squares on the wall) are caused by large intensity variations between neighboring pixels, not by depth cues. (b) An HDOF image using this erroneous variance map reflects the errors in, for example, the wall area, while correctly showing dimming effects in the left and right toys. (Light field images courtesy of the Stanford light field database).

While the proposed Invert and Compare (IC) algorithm achieves a smooth gradient over almost the entire depth range, local pixel-scale ringing artifacts (i.e., Figure 5(g,h) middle) may be observed in the IC method's variance map. There are weak artifacts in the focused region of the regular image (Figure 5(g) left), caused by the Black Matrix (BM) pattern between projector pixels. These BW artifacts appear in both the IC ÒsumÓ image as well as its variance map (Figure 5(g,h) middle). The mesh-like artifact is further enhanced in regions of the IC variance map with slight defocus, due to subtraction of two blurred checkerboard spots that have not spread completely into neighboring spots of opposite value. This defocus range depends upon camera parameters and object distance, but is on the order of a centimeter. A simple mathematical morphology closing operation on the order of 1 to 3 pixels applied to the whole variance map compensates nicely for this effect (Figure 5(g,h) right). This operation can degrade the accuracy of depth cues on the order of 1 to 3 pixels. However, considering the designs aim to enhance visualization of DOF, such a loss of accuracy is acceptable. An example of this slight loss in high-frequency detail is observable by comparing Figure 5(g) and 5(h) right, which is the same region of the red box in Figure 9(a,b)

To summarize the procedure for the two-shot method:

(1) capture a checkerboard pattern image;
(2) invert the pattern and capture a second image;
(3) generate a variance map from the normed absolute difference of the images;

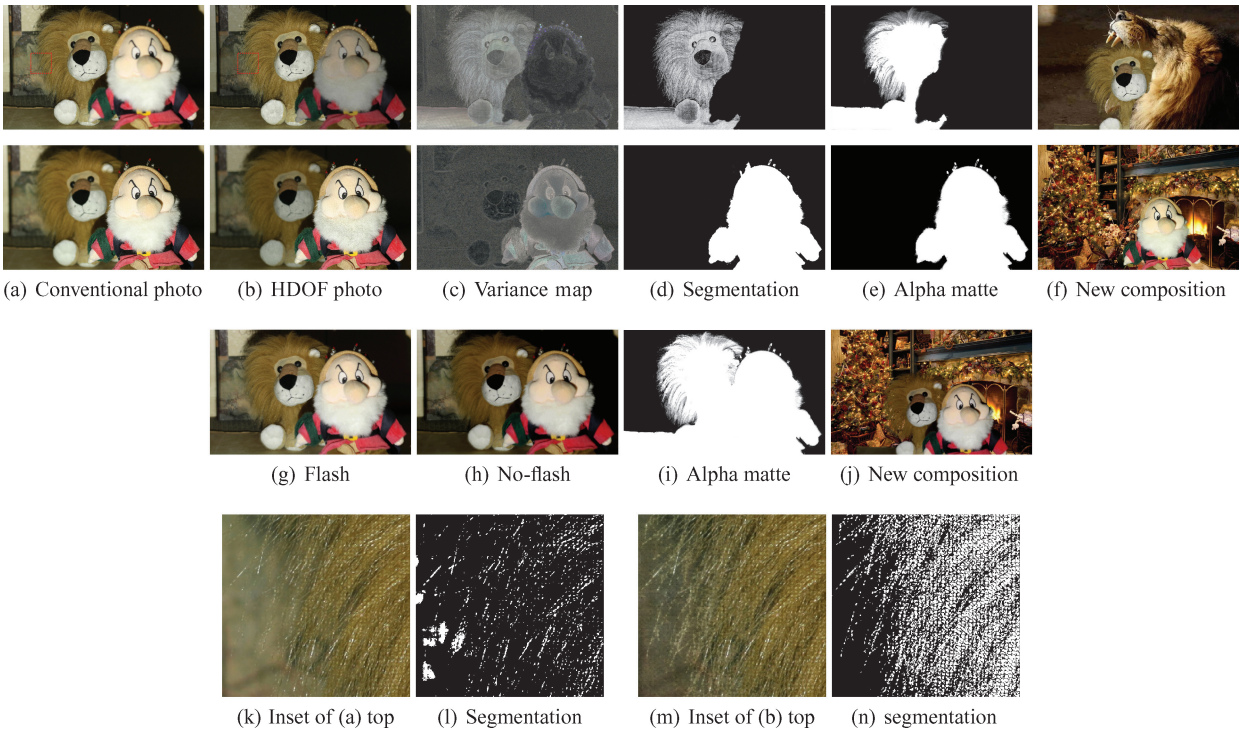


Fig. 9. (a)-(f) The two-shot IC HDOF method can be applied to shape segmentation and depth-dependent matting for objects at different focal planes. (g)–(j) In comparison to flash matting, which is also a two-shot matting technique, the HDOF method can segment two objects at slightly different depths. (k)–(n) The HDOF technique (n) enhances visibility and segmentation of high-frequency details compared to segmentation of a conventional photo (l).

(4) perform a closing operation for the whole variance map to compensate both BM and meshing artifacts;
(5) multiply the variance map with the two images' sum.

4. EXPERIMENTAL RESULTS

The imaging system used to demonstrate focus-dependent intensity gradients consists of a board-level 4008 × 2672 pixel color Lumenera CCD and a Sigma f/2.8 zoom lens kept at a fixed setting. We use a Mitsubishi PK10 Pocket Projector with 800 × 600 resolution for our grid illumination source. The projector was placed next to the camera and focused roughly on the camera's plane of focus. The projector's depth of field does not need to be infinite but should be larger than the scene depth of interest. Pocket projectors as well as laser projectors have a very large depth of field. The grid of point sources remained in-focus for the scene depths and object distances tested (20–100cm).

4.1 Variable Depth of Field Demonstration

A simple calibration experiment verified the analysis presented in Section 2 and demonstrated the concept of a variable intensity-dependent DOF using our secondary processing methods. The object we imaged was a uniformly lit planar Lambertian object tilted with respect to the z-axis, as illustrated in Figure 5. Testing the multi-shot method (SM) first, we projected grid patterns with different duty cycles and captured multiple images to search for pixel maxima. Four different patterns were shifted, with duty cycles of $1/n^2$ (i.e., requiring n^2 images), with n ranging from 2 to 5. Each projection point was created from a single projector pixel.

A result from applying the algorithm in Eq. (6) to these sets of images is displayed in the lower right of Figure 5. In addition to testing the multishot method with this tilted object, we also projected and shifted a checkerboard pattern and captured two images to test the two-shot IC method. Note that its intensity-dependent DOF is much narrower and sharper than the result of other patterns. This can be attributed to the pixel weighting based on an absolute difference. Such sharp results requiring only two images make this approach quite promising.

4.2 Comparison of Three Methods

A duty cycle of 1/9 was found to be the projected pattern period that led to an optimal trade-off between sensor resolution and intensity drop for both the single-shot and multiple-shot methods. Figure 6 compares conventional photos and HDOF photos generated using the single- and multishot methods in reduced and full resolution, respectively. Note that the resolution reduction using the single-shot method introduces pixelation and artifacts due to imperfect sampling of the original image, visible in the inset. Example photographs captured with our secondary methods of image capture and processing are in Figure 7. It is clear that both approaches highlight the focal plane for a varying range of lens settings, and can significantly dim close objects that would otherwise reflect more light to the lens area than distant objects. Furthermore, these examples show that our two secondary methods can handle a range of scene complexity and a limited degree of reflectivity.

4.3 Segmentation and Matting with Two-Shot Method

Our two-shot method offers a unique feature useful for detailed shape segmentation and automatic matting, shown in Figure 9. As noted in Section 3.3, comparing captured checkerboard pattern images yields a depth-dependent variance map (Figure 9(c)).

(a) Conventional (b) HDOF (c) inset

Fig. 10. Limitations of the HDOF technique. (a) A conventional photo focused on the foreground subject (top) and background subject (bottom). (b) The corresponding HDOF photo created with the IC method. (c) The inset of the red box in (b), showing bright artifacts around the subject's, hair due to the movement of hair between two continuous shots (top), as well as background noise caused by strong ambient light washing out the projected pattern on the wall (bottom).

A segmentation image is generated by simple thresholding and blob analysis of this variance map (Figure 9(d)). The segmentation image is used as a trimap for an automatic application of matting (Figure 9(e) and Figure 9(f)). The comparison of Figure 9(l) and Figure 9(n) demonstrates the precision of our segmentation process with similarly colored foreground and background objects (Figure 9(k) and Figure 9(m)). A previous matting technique called flash matting [Sun et al. 2006] also uses two images, captured with and without a flash. Whereas flash matting works only for foreground objects that are sufficiently distant from the background (Figure 9(g)–(j)), our method works for multiple objects at a range of focal planes, selected simply by refocusing the lens (Figure 9(e) top and bottom). However, comparing the lion hair in Figure 9(e) and Figure 9(i), it is clear that the flash/no-flash technique emphasizes fine structures with more precision than the proposed HDOF technique. Similarly, Joshi et al. [2006] presented a system for natural video matting using a camera array. Unlike our setup, they used 8 cameras to generate a variance map, which fails to offer segmentation for a scene with a uniform background. Figure 8(b) provides a demonstration of this concept, where a test HDOF image is generated using different focal stack images from a camera array. In the variance map (Figure 8(c)), errors are clear in the background due to sharp color changes amongst neighboring pixels. Our active-illumination-based HDOF system provides a robust variance map for segmentation and matting for any scene within a compact setup.

4.4 Limitations

Similar to other active illumination systems, ambient light can overpower the projected patterns and decrease SNR (Figure 10(c) bottom), causing our setup to work best in settings where the pattern is clearly visible. For example, our experiments were performed indoors in normally lit rooms. Additionally, emissive objects, specularities, and complex inter-reflections can present difficulties. The requirement of a wide-DOF projector may limit this technique to certain environments. Additional limitations are specific to each setup. The single-shot method reduces the original sensor resolution by a factor of 9. The other two methods rely on two or more images, thus working best with static scenes (Figure 10(c) top).

5. CONCLUSION

We have described a new method of photography, using an illumination pattern which can present the defocused content of a scene in the form of an intensity gradient. The primary design that achieves this in a single image does so at the expense of sensor resolution. Two secondary designs make up for this loss through two or more images and computation, producing full resolution results. The clear benefits of this photographic technique to artistic composition and contrast enhancement were demonstrated experimentally. Future work will look at extending the use of this camera to achieve optical foveation, where only the salient (focused) features of a scene are captured and defocused content is blocked. This has the potential to decrease bandwidth on a high-throughput system, or could possibly be used as an aid in image compression. Our design could also benefit realms parallel to microscopy, such as in underwater photography, in which turbid 3D media affects overall photographic resolution. Finally, the HDOF photograph's potential to assist with object segmentation can be guided towards object identification.

ACKNOWLEDGMENTS

We thank Fredo Durand and Ankit Mohan for constructive comments, and Manuel M. Oliveira and Yasuhiro Mukaigawa for thorough article revisions. Finally, thanks to Sharmeen Browarek for her help with narrating a successful summary video.

REFERENCES

BAE, S. AND DURAND, F. 2007. Defocus magnification. *Comput. Graph. Forum 26,* 3.

BORN, M. AND WOLF, E. 1970. *Principles of Optics*. Pergamon Press.

DAVIS, J., NEHAB, D., RAMAMOOTHI, R., AND RUSINKIEWICZ, S. 2005. Spacetime stereo : A unifying framework for depth from triangulation. *IEEE Trans. Patt. Anal. Mach. Intell. 27,* 2.

EISNER, M., LINDLEIN, N., AND SCHWIDER, J. 1998. Confocal microscopy with a refractive microlens-pinhole array. *Optics Lett. 23,* 10.

HASINOFF, S. AND KUTULAKOS, K. 2006. Confocal stereo. In *Proceedings of the 9th European Conference on Computer Vision (ECCV)*, 259–268.

HEINTZMANN, R., HANLEY, Q. S., ARNDT-JOVIN, D., AND JOVIN, T. M. 2001. A dual path programmable array microscope (pam):simultaneous acquisition of conjugate and non-conjugate images. *J. Microscopy 204*, 119–137.

JOSHI, N., MATUSIK, W., AND AVIDAN, S. 2006. Natural video matting using camera arrays. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques. ACM SIGGRAPH.* 779–786.

KIM, J., LANMAN, D., MUKAIGAWA, Y., AND RASKAR, R. 2010. Descattering transmission via angular filtering. In *Proceedings of the European Conference on Computer Vision (ECCV'10)*. Lecture Notes in Computer Science, vol. 6311. Springer, 86–99.

LAI, S., FU, C., AND CHANG, S. 1992. A generalized depth estimation algorithm with a single image. *IEEE Trans. Patt. Anal. Mach. Intell. 14,* 4, 405–411.

LEMOIGNE, J. AND WAXMAN, A. 1988. Structured light patterns for robot mobility. *IEEE J. Robot. Autom. 4,* 5, 541–548.

LEVOY, M., CHEN, B., VAISH, V., HOROWITZ, M., MCDOWALL, I., AND BOLAS, M. 2004. Synthetic aperture confocal imaging. In *Proceedings of the Conference on Computer Graphics and Interactive Techniques. ACM SIGGRAPH*.

LEVOY, M., NG, R., ADAMS, A., FOOTER, M., AND HOROWITZ, M. 2006. Light field microscopy. *ACM Trans. Graph. 22,* 2.

LEVOY, M., ZHANG, Z., AND MCDOWALL, I. 2009. Recording and controlling the 4d light field in a microscope. *J. Microscopy 235.*

MAAS, H. 1992. Robust automatic surface reconstruction with structured light. *Int. Arch. Photogram. Remote Sens. 29,* B5.

MITIC, J., ANHUT, T., SEROV, A., AND LASSER, T. 2003. Real-Time optically sectioned wide-field microscopy employing structured light illumination and a cmos detector. *Proc. SPIE 4964*.

MORENO-NOGUER, F., BELHUMEUR, P. N., AND NAYAR, S. K. 2007. Active refocusing of images and videos. In *Proceedings of the Conference on Computer Graphics and Interaction Techniques ACM SIGGRAPH* 2007.

MOUADDABI, E., BATILE, J., AND SALVI, J. 1997. Recent progress in structured light in order to solve the correspondence problem in stereovision. In *Proceedings of the IEEE International Conference on Robotics and Automation. (ICRA)*.

NAYAR, S., KRICHNAN, G., GROSSBERG, M., AND RASKAR, R. 2006. Fast separation of direct and global components of a scene using high frequency illumination. *ACM Trans. Graph. 25,* 3, 935–943.

NG, R., LEVOY, M., BREDIF, M., DUVAL, M., HOROWITZ, G., AND HANRAHAN, P. 2004. Light field photography with a hand-held plenoptic camera. Tech. rep, Stanford University.

SALVI, J., PAGES, J., AND BATLLE, J. 2004. Pattern codification strategies in structured light systems. *Patt. Recogn. 37*, 827–849.

SCHECHNER, Y., KIRYATI, N., AND BASRI, R. 2000. Separation of transparent layers using focus. *Int. J. Comput. Vis. 39,* 1, 25–39.

SHRIKHANDE, N. AND STOCKMAN, G. 1989. Surface orientation from a projection grid. *IEEE Trans. Pattern Anal. Mach. Intell. 11,* 6, 650–655.

SUN, J., KANG, S. B., , AND SHUM, H. Y. 2006. Flash matting. In *Proceedings of the International Conference on Computer Graphics and Interactive Techniques. ACM SIGGRAPH*. 361–366.

TIZIANI, H. AND UHDE, H. 1994. Three-Dimensional analysis by a microlens-array confocal arrangement. *Appl. Optics 33*, 567–572.

WANG, Y., MITICHE, A., AND AGGARWAL, J. 1987. Computation of surface orientation and structure of objects using grid coding. *IEEE Trans. Pattern Anal. Mach. Intell. 9*, 129–137.

WATANABE, M. AND NAYAR, S. 1998. Rational filters for passive depth from defocus. *Int. J. Comput. Vis. 27,* 3, 203–225.

WILL, P. AND PENNINGTON, K. 1971. Grid coding: A preprocessing technique for robot and machine vision. *Artif. Intell. 2*, 319–329.

WILSON, T., JUSKAITIS, R., NEIL, M., AND KOZUBEK, M. 1996. Confocal microscopy by aperture correlation. *Optics Lett. 21,* 3.

ZHANG, L. AND NAYAR, S. 2006. Projection defocus analysis for scene capture and image display. *ACM Trans. Graph. 25,* 3, 907–915.

Received August 2010; revised January 2011; accepted March 2011