Sensing Increased Image Resolution Using Aperture Masks

Ankit Mohan, Xiang Huang, Jack Tumblin
EECS Department, Northwestern University
http://www.cs.northwestern.edu/~amohan

Ramesh Raskar*
Mitsubishi Electric Research Laboratories
http://web.media.mit.edu/~raskar

Abstract

We present a technique to construct increased-resolution images from multiple photos taken without moving the camera or the sensor. Like other super-resolution techniques, we capture and merge multiple images, but instead of moving the camera sensor by sub-pixel distances for each image, we change masks in the lens aperture and slightly de-focus the lens. The resulting capture system is simpler, and tolerates modest mask registration errors well. We present a theoretical analysis of the camera and image merging method, show both simulated results and actual results from a crudely modified consumer camera, and compare its results to robust 'blind' methods that rely on uncontrolled camera displacements.

1. Introduction

High-quality lens systems are expensive and difficult to design, and in spite of ever-increasing sensor resolution, conventional digital cameras don't really capture all the sharpness and details of a real world scene. While 'super-resolution' techniques attempt to recapture some of these lost details, only a few high-end camera designs apply them, because the camera must precisely adjust its image sensor position in sub-pixel increments for each of a series of k photos.

This paper describes a new super-resolution technique that also requires a series of k photos, but eliminates the need to move the image sensor by sub-pixel distances. Instead, for each of the k photos we use a relatively poor quality lens (or slightly de-focus a high quality lens), and place a different mask at or near the lens' limiting aperture. Each mask is a coarse grid of transparent and opaque squares, and the set of masks partitions the lens aperture into k pieces. We obtain resolution improvements comparable to existing k-photo methods, but the implementation is much easier and does not require a cumbersome image registration process as the optical axis, lens and sensor remain fixed. Instead of sub-pixel sensor displacements (sub-micrometer) using high-precision actuators, masks that subdivide the aperture cover and reveal large areas (tens of millimeters). As our method only requires changing masks inside or on the lens, it may prove suitable for retrofitting on a broad range of existing imaging systems from microscopes to telescopes. Figure 1 shows a low-cost, commercially-

Figure 1. (Top) Modified filter-holder on a Canon Digital Rebel XT SLR camera accepts different masks slid in front of the lens. (Bottom) Four masks mounted on the camera to obtain 2×2 resolution enhancement.

*now at the Media Lab, MIT.

978-1-4244-2243-2/08/$25.00 ©2008 IEEE

available camera, modified to accept these masks near the aperture plane, and even this simple experimental setup shows worthwhile resolution improvements. The resulting capture system is simpler, cheaper, and tolerates modest mask registration errors well. Better designs might fit selectable masks inside the lens, or use LCD or DMD devices for arbitrary programmable apertures [22].

1.1. Contributions

By exploiting intentional blur and carefully designed aperture masks, we demonstrate a new method for achieving increased resolution using multiple photos.

- We achieve controllable sub-pixel image shifts via intentional blurring and changing aperture patterns. This is similar to Adelson and Wang's [1] use of eccentric apertures for acquiring the light field of a scene.
- We link the degree of defocus, mask pattern and size, and resolution improvement it can provide.
- We develop a reconstruction process to make one high resolution image from several low resolution photos taken through different aperture masks.

Unlike several other methods, our technique only provides enhanced resolution for the scene elements within the camera's depth of field. Out-of-focus elements of the scene are not enhanced, and the technique works best for fronto-parallel scenes. We also require a constant known blur on the sensor, which is easy to achieve using optical and mechanical methods. Additionally, our technique shares the diffraction limit of the lens system, and we cannot improve resolution of a diffraction limited system. Since we are using reduced size apertures, our techniques are ideal when the system is at least an order of magnitude from the diffraction limit, and may prove most helpful with low-quality lenses where the lens point spread function (PSF) is approximately equal to the pixel size.

1.2. Related Work

Super resolution refers to techniques that improve image resolution by combining multiple low-resolution images to recover higher spatial frequency components lost to undersampling. Numerous algorithms and techniques have been proposed [20, 13, 9, 10, 14, 4], which first estimate the relative motion between the camera and the scene, and then register all images to a reference frame. Then they fused the images, usually by interleaving filtered pixels, to obtain a high resolution image. Keren et al [12], and Vandewalle et al [21] used randomized or 'jittered' sensor positions that they estimated using sub-pixel image registration. Komatsu et. al. [15] integrated images taken by multiple cameras with different pixel apertures to get a high resolution image. Joshi et. al. [11] merged images taken at different zoom levels, and Rajan et al. [17] investigated the use of blur, shading, and defocus for achieving super-resolution of an image and its depth-map. Most authors also applied modest forms of de-convolution to boost the image's high spatial frequency components that were reduced by the 'box' filter. Park et al. [16] and the book by Chaudhuri [2] provide a unified survey and explanation of many current methods.

The idea of putting a mask in the aperture of an optical system has been well explored in the astronomy and scientific imaging [19]. Refined "coded-aperture" methods using MURA arrays [7] enabled lens-less gamma-ray imaging systems for distant stars. Donoho [3] and Fergus et al. [6] use a random aperture configuration to acquire random projections of a digital image. Farid [5] exploit defocus using masks in the aperture for range estimation. Our goal is achieve a combination of defocus blur and masks to achieve an orderly sequence of fractional-pixel displacements of the image.

2. Mask-based Super-resolution Method

We gather k nearly-identical photos and merge them. Counterintuitively, we intentionally defocus the camera lens by a tiny amount, and place different masks in the limiting aperture for each photo to cause sub-pixel image translation. The photo merging step interleaves photo pixels to match their displacements, and then de-convolves this spatially-super-sampled image to counteract the blur caused by the lens, mask, and the pixel sensor area as well.

Our method resembles early super-resolution methods that translated the image sensor by precise, fractional-pixel amounts for each photo, so that the interleaved pixel sampling grids form a new uniform grid at a multiple of the original sampling frequency. Such precise translation is problematic, as typical pixel sensors span only a few micrometers. Some systems employ servos with active feedback, and others discard the goals of uniform sampling altogether. Systems such as [12, 21] ask users to intentionally perturb the entire camera by modest but unknown ('blind') amounts, estimate sub-pixel offset and rotation, and then merge samples at their presumed locations to reconstruct super-resolution output. While results often look good, their quality is unpredictable for small numbers of photos. Our system is simpler, and as it forms precisely-translated images. In addition, its output is suitable as high-quality input for 'blind' alignment software, which can detect and correct minor positional errors caused by imprecise mask size, position, and de-focus amounts.

2.1. How Masks Translate the Sensor's Image

Figure 2 illustrates how changing aperture masks can change the translation amounts for an image focussed

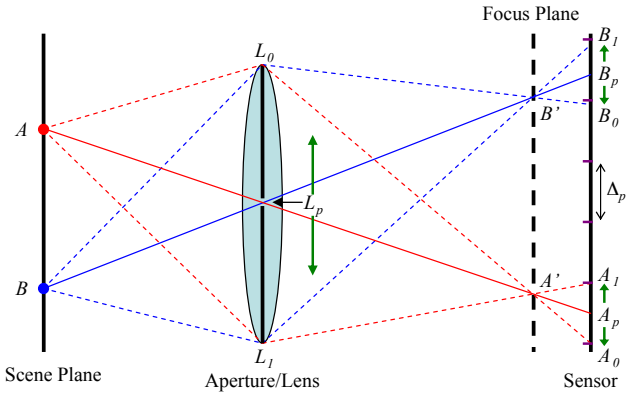


Figure 2. Our camera setup with moving pinhole apertures. The sensor is moved slightly from the plane of focus so that the point spread function is exactly one pixel. As the pinhole (L_p) moves across the lens from L_0 to L_1, the image of scene points A and B move from A_0 to A_1 and B_0 to B_1. The effect is the same as when the sensor is shifted by a fraction of the pixel size.

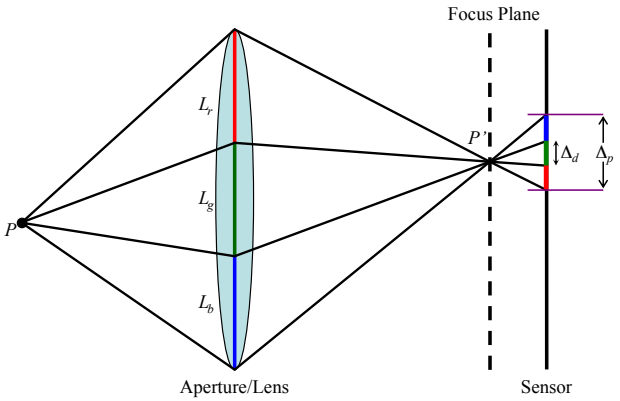


Figure 3. Our camera setup for finite sized apertures. Three different partial-aperture masks ($k = 3$) create three different, slightly blurred images on the camera sensor, each one translated by $1/k = 1/3$ pixel spacing. Correct lens de-focus maps scene point P through three aperture regions (L_r, L_g, or L_b) to exactly one sensor-pixel area: $k\Delta_d = \Delta_a = \Delta_p$.

slightly above or below the sensor itself. We will describe our method for the plane shown in the drawing, a 1D slice of a 2D imaging system. We assume the photographed scene is planar and perpendicular to the optical axis of the camera lens. Also, the sensor fill rate is 100%, each pixel has width Δ_p.

The lens refracts all the rays that pass through it, and expanding cones of rays (dotted lines) from each scene point A and B converge to points A' and B' respectively on the focus plane. On the sensor, the rays from point A spread to cover $(A'_0..A'_1)$ and the point B rays spread to cover $(B'_0..B'_1)$. This 'circle of confusion', or the 'point spread function' (PSF) is simply a scaled version of the lens aperture L. Placing a single movable pinhole aperture L_p in the lens blocks almost all the light, and only one ray from each scene point arrives at the sensor. As the pinhole moves from L_0 to L_1 the image of scene points A, B move from A_0, B_0 to A_1, B_1 on the sensor. If we choose our defocus amount so that A_0 to A_1 spans exactly one pixel spacing Δ_p, then equally spaced pinhole locations will translate the image by precise fractions of a pixel. The de-focus amount and k aperture masks work together to achieve precise fractional displacements we need for super-resolution. If we know the exact amount of de-focus, we can find the best aperture masks, or any de-focus amount that forms a PSF larger than Δ_p lets us reduce that PSF to a series of k apertures offset by Δ_p/k.

2.2. Blur from Square Mask Apertures

In theory, pinhole apertures produce perfectly focussed images, but in practice our masks need apertures as large as possible to minimize light loss, to minimize sensor noise, and to avoid diffraction effects. However, larger apertures blur the sensor image, which is the convolution of the translated sharp image of an ideal pinhole aperture and the scaled shape of the aperture itself (PSF). As shown in Figure 3, the PSF of an ideal diffraction-free lens with an open *square*-shaped aperture is a 'box' function of size Δ_a. For our slightly-defocussed camera with a fully open aperture, the PSF size at the sensor is equal to the pixel-to-pixel spacing Δ_p. For a partially-blocked aperture, the PSF size Δ_d is even smaller.

While each of the k photos has the same number of pixels, each of the k photos are different as long as the image displacement is distinct and less than one pixel. The aperture mask effectively sets the shape of the out-of-focus PSF of the camera's lens. Our goal is to compute the unknown 1D scene signal, $s(x)$ (see Figure 4). We denote the lens PSF at the sensor plane corresponding to the i^{th} aperture mask (L_i) by $l_i(x)$. The image signal arriving at the sensor plane is given by the convolution: $s(x) * l_i(x)$, and each pixel's square light sensing area acts as a 'box' filter $p(x)$ that further limits image bandwidth and helps suppress aliasing artifacts. The continuous image, $f_i(x)$, that we sample at the sensor plane is $f_i(x) = s(x) * l_i(x) * p(x)$. The sensor collects a digital image $f_i[x]$, by sampling $f_i(x)$ at pixel locations:

$$f_i[x] = (s(x) * l_i(x) * p(x)) \cdot \text{Ш}(x/\Delta_p)/\Delta_p, \quad (1)$$

where Ш is the impulse train: $\text{Ш}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n)$, and Δ_p is the pixel-to-pixel spacing.

If we photograph through a mask, we can describe the i^{th} aperture mask as the convolution of a translated impulse function (pinhole aperture) with a box function, $l_i(x) =$

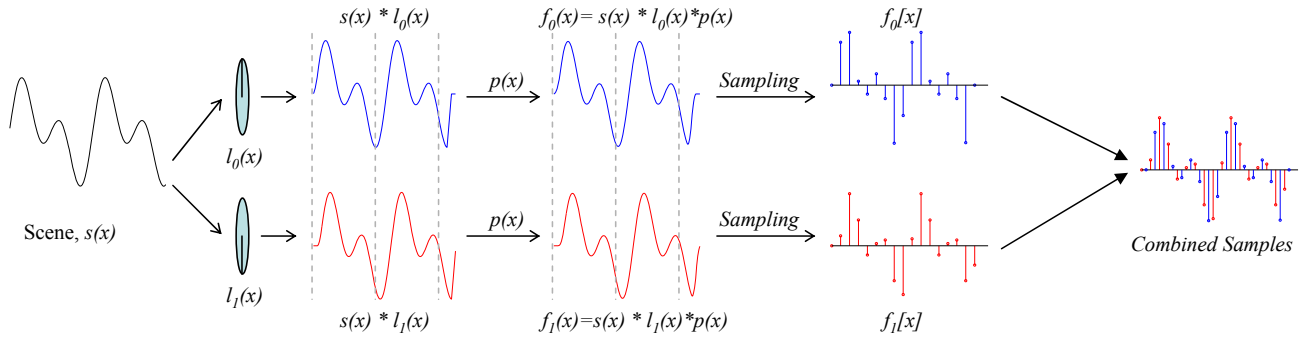


Figure 4. Enhancing resolution by 2× for a 1D signal $s(x)$: We start with capturing two photos of the scene with two orthogonal masks in front of the lens, $L_0(x)$ and $L_1(x)$. Slightly out of focus, the image at the sensor is convolved with by a scaled version of the aperture mask (and the lens PSF). Both signals arrive at the sensor *shifted by half the pixel spacing*. Each pixel's sensor applies additional blur $p(x)$, that aids antialiasing. As the two sensed images $f_i[x]$ differ by a half-pixel-spacing shift, pixel interleaving provides some higher-frequency details. De-convolving by $p(x)$ and $l_i(x)$ (not shown) further improves results.

$\delta(x - x_i^p) * d(x)$, where $d(x)$ is the box function with PSF width Δ_d on the sensor plane, and x_i^p is the position of that pinhole aperture. From Equation 1 we have,

$$f_i[x] = (s(x) * p(x) * d(x) * \delta(x - x_i^p)) \cdot \text{Ш}(x/\Delta_p)/\Delta_p.$$

Substituting $s_d(x) = s(x) * d(x)$, and shifting the axis, we can write,

$$f_i[x] = (s_d(x) * p(x)) \cdot \text{Ш}\Big(\frac{x - x_i^p}{\Delta_p}\Big)/\Delta_p \qquad (2)$$

These equations lead to four important observations about our system. First, we need to know the blur size of the camera with an open aperture, and this should preferably be equal to the pixel-to-pixel spacing ($\Delta_a \approx \Delta_p$). Second, we must design our aperture masks L to impose a series of k unique displacements x_i^p that subdivide the pixel spacing Δ_p into uniform fractional amounts. Third, we must recognize that traditional sensor-shifting super-resolution imposes no defocus blur, but both are subjected to the 'box' filter convolution imposed by each pixel's light-sensing area. Finally, for a resolution enhancement factor k, $\Delta_d = \frac{1}{k}\Delta_p$; the mask-imposed blur is $\frac{1}{k}$ times smaller than the area covered by a single pixel sensor. However, its fractional size will exactly match the pixel spacing in the estimated high-resolution image we assemble from these k shifted photos.

3. Reconstruction

Most translation-based super-resolution methods apply a three step reconstruction process: registration, interpolation, and restoration. We simplify the registration step as we carefully control the blur and use known aperture masks; thus we *know* the exact PSF and image shifts for the k photos. We represent the image formation process in the discrete domain using a defocus or 'blur' matrix B and a decimation matrix D. This representation reduces the problem of estimating an increased resolution image to one of inverting k equations that describe the unknown scene intensities s. We represent the n pixel image captured with the i^{th} aperture mask as a vector, $\mathbf{f_i}$. The unknown scene information, a 1D vector of length nk, is $\mathbf{s}$, where k is the desired resolution enhancement factor. The image formation process is then,

$$\mathbf{f_i} = \mathbf{D} \cdot \mathbf{B_i} \cdot \mathbf{s}, \qquad (3)$$

The defocus or 'blur' matrix $\mathbf{B_i}$ describes how i^{th} aperture mask and out-of-focus lens modifies the ideal image s (convolution of the scene with blur $l_i(x)$ in Equation 1). The decimation matrix represents the effect of the antialiasing filter ($p(x)$) due to the finite pixel sensor size, followed by sampling and reduction in resolution by a factor of k.

As described earlier, the blur size due to the open aperture (Δ_a) is equal to the sensor pixel size. In order to achieve a resolution enhancement factor of k, the blur size due to any partial aperture (Δ_d) is equal to $\frac{1}{k}$ of the sensor pixel size. However, aperture masks can hold any linearly independent combination of partial apertures we wish–not just a single $\frac{1}{k}$ opening.

Consider a k element 1D aperture mask, $\mathbf{m_i} = \{m_i^0, \ldots, m_i^{k-1}\}$, where each element indicates transparency ($0 \le m_i^j \le 1$) of the corresponding mask cell. When the element is opaque, $m_i^j = 0$. The general defocus matrix for the i^{th} photo has the form,

$$\mathbf{B_i} = \begin{pmatrix} m_i^0 & \cdots & m_i^{k-1} & & \\ & \ddots & \ddots & \ddots & \\ & & m_i^0 & \cdots & m_i^{k-1} \end{pmatrix}.$$

The dimensions of matrix $\mathbf{B_i}$ are $(nk + k - 1) \times nk$. In the specific case where the mask is made up of a single transparent element, i.e., $\mathbf{m_i} = \{0, \ldots, 0, 1, 0, \ldots, 0\}$, each $\mathbf{m_i}$

is orthogonal to each other. In this case, $m_i^j = 1$ for $i = j$, and 0 otherwise. However, these simple masks allow only a small amount of light to pass through the partial aperture. By using Hadamard codes instead [8] we can improve the noise performance, because binary codes from its S-matrix block only about half the light entering the lens.

The decimation matrix includes both the reduced sampling rate from one image and the low-pass filter imposed by the area of each pixel sensor:

$$\mathbf{D} = \begin{pmatrix} \ddots & \overbrace{1 \cdots 1}^{k \text{ ones}} & & \\ & & 1 \cdots 1 & \\ & & & \ddots \end{pmatrix}.$$

The dimensions of matrix $\mathbf{D}$ are $n \times (nk + k - 1)$.

The simplified registration step, in our case, involves determining the defocus matrices, $\mathbf{B_i}$, based on the known mask patterns and their PSF. This is trivial in our case if the blur size of the open aperture exactly equals the physical pixel size. Inverting the composite matrix created by stacking the equations for each image observation, $\mathbf{f_i}$, gives us nk equations and as many unknowns.

To summarize, the technique involves the following steps.

- Photograph the plane with scene signal s and image onto the n-pixel sensor so that the defocus blur PSF is exactly one pixel for an open aperture.
- Take k photos, each with a different partial aperture mask L_i so that the PSF is $l_i(x)$, and measure the discretized n-value sensor signal $f_i[x]$.
- Solve Equation 3 in the discrete domain to compute nk values of signal s.

4. Extending to Arbitrary Blur size

So far we have discussed the case where the image blur is exactly equal to the size of one sensor pixel. While this makes the super-resolution process simpler, it is not a requirement for our method, and in this section we generalize our method to work for an arbitrary but known blur size.

Consider the case where the image blur is uniform across the image (fronto-parallel scene uniformly out of focus), and the size of the circle of confusion is $\Delta_a = b \times \Delta_p$ (as shown in Figure 5). Here b is the blur size in the number of pixels (it does not have to be a whole number). Just like the single pixel blur case (Figure 2), we want to shift the image by Δ_p/k, where Δ_p is the sensor pixel size, and k is the desired super-resolution factor. We can still achieve this by positioning a pinhole at k distinct positions in the aperture plane and capturing k photos. These pinhole positions are

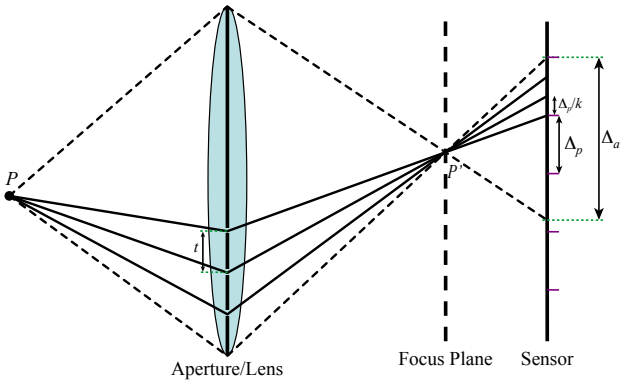


Figure 5. Camera setup for an arbitrary blur size.

equally spaced by $t = \frac{a}{kb}$ between any nearest pair, were a is the size of the open aperture. For a blur of exactly one sensor pixel, this reduces to $t = a/k$ as discussed in Section 2.1.

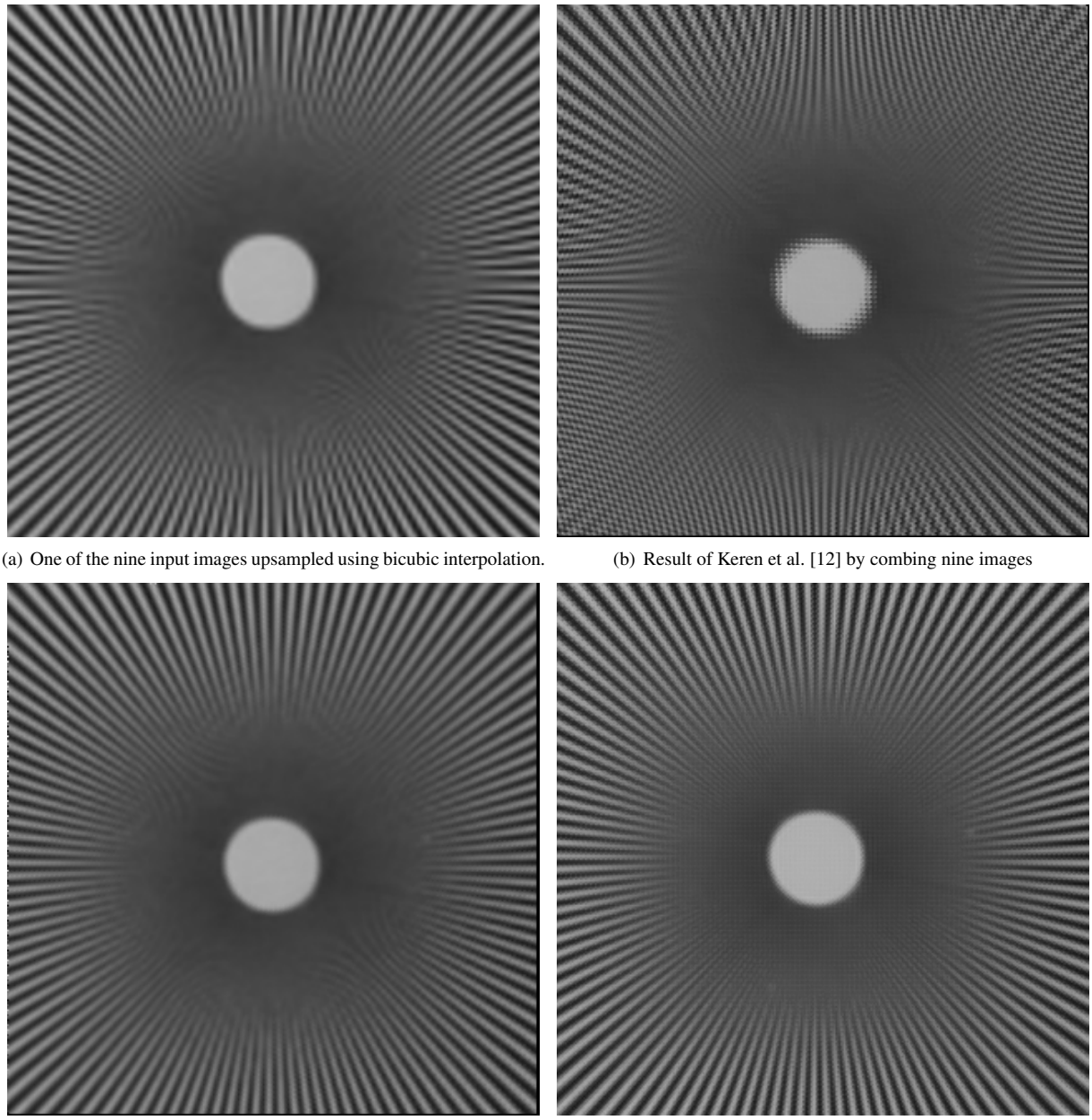
We can use finite sized apertures instead of pinholes, just like the single pixel blur case as shown in Section 2.2. By using a square aperture with a size $d = \frac{a}{kb}$, and moving it by the same distance ($t = \frac{a}{kb}$), we blur the captured image by Δ_p/k. Just like the unit pixel blur case, this is the same size as the pixel spacing in the reconstructed high-resolution photo. This is certainly better than using a pinhole, but it still throws away a large amount of the light otherwise captured by the fully open aperture, a.

Unfortunately the aperture mask size depends on the exact blur size. This, and the fact that we throw away increasingly large parts of the light as blur size grows makes arbitrarily large blur sizes undesirable. Instead, we advocate the use of optical and mechanical means to ensure a unit pixel blur. Adjusting the auto-focus system of a camera to cause a fixed blur in an image should be fairly straightforward given access to the camera firmware. We believe this task is best left to the camera manufacturers. We discuss our method of simulating a unit-pixel blur in Section 5.

5. Results

We used a 8-megapixel Canon Digital Rebel XT with a Canon 50mm/1.8 lens for building our prototype. We used a Cokin "A" series filter holder to slide and hold the masks in front of the lens (as shown in Figure 1). The masks were printed on transparent film using an inkjet printer. With some care, the masks slide into and out of the filter holder without moving the camera. The filter holder also provides nice alignment tabs to ensure the masks are in the right position.

We placed the mask outside the lens, and not in its aperture plane. While this is not the optimal position for the mask, the external mount lets us change masks much more



(a) One of the nine input images upsampled using bicubic interpolation.

(b) Result of Keren et al. [12] by combing nine images

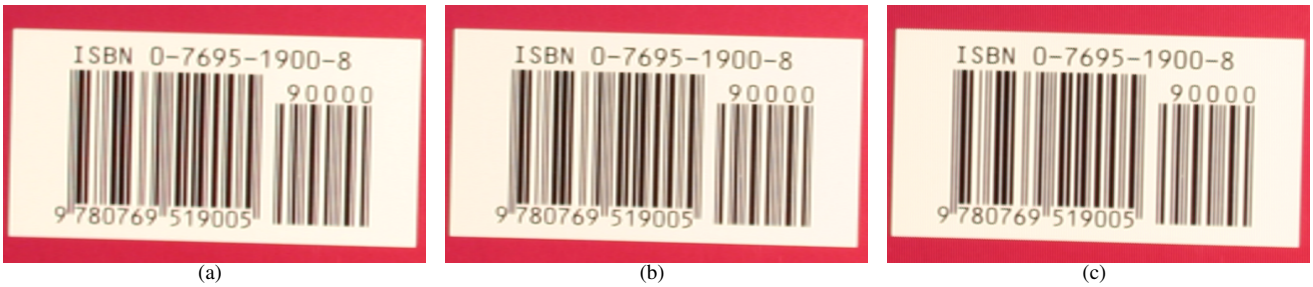
(c) Result of Vandewalle et al. [21] by combing nine images

(d) Our result by combining the nine images

Figure 6. 3x resolution enhancement in 2D. (a) shows one of the input images upsampled using bicubic interpolation; (b) and (c) are obtained from super-resolution algorithms that attempt to estimate camera motion in a number of photos and merge the information. Our result looks cleaner, has more detail, and reduced aliasing and moire patterns (d).

easily. Carefully chosen mask sizes ensure that the mask acts as the limiting aperture of the optical system. If this is not the case, the mask can cause vignetting, and result in an undesirable spatially varying PSF on the image plane.

For all the results, we introduce an arbitrary blur by slightly defocusing the camera lens. We estimate the blur size by capturing two extra photos, each with a mask in the camera aperture plane. Each mask is completely opaque other than a small opening at opposite corners. This gives two photos shifted by a distance equal to the size of the circle of confusion. We use the frequency domain registration approach of Vandewalle et al [21] to register these two photos with sub-pixel accuracy. This blur is usually 1-2 pixels. To simulate a unit-pixel blur, we down-sample the input images so that the effective blur size equals the down-sampled pixel spacing. Alternatively, we could use a fixed



(a) (b) (c)

Figure 7. 4x resolution enhancement in 1D along the horizontal direction. (a) and (b) are two of the four input images upsampled horizontally using bicubic interpolation. (c) is our result with enhanced sharpness, and reduced aliasing of the vertical bars.

(a) (b) (c)

Figure 8. 4x resolution enhancement in 1D along the horizontal direction. (a) and (b) are two of the four input images upsampled horizontally using bicubic interpolation. (c) is our result by combining the four images.

set of masks and set the matching amount of lens de-focus, perhaps by adjusting the manual-focus ring. For our lens, these adjustments were too tiny to be practical. While neither of these offer a truly practical solution, we believe they adequately demonstrate the effectiveness of our technique.

Figures 6, 7, and 8 show some of the results of our system. Please note that some of the differences are quite subtle and may not be visible on a printed version. Please view the supplemental material for more results. Figure 6 show the results obtained by image registration and merging algorithms of [21, 12]. Since the image shift is non-uniform and may be inaccurately estimated, these methods are not as effective at enhancing resolution and minimizing aliasing as our approach.

6. Discussions and Future Work

Limitations: Perhaps the biggest limitation of our technique as presented is that we require the scene to be frontoparallel, and uniformly out of focus by a known amount. While this might seem as a big limitation, we argue that most often enhanced resolution is required only in the parts of the image that are in focus, and out-of-focus parts would gain very little from super-resolution. Also, we believe that achieving a fixed known blur is just as easy as optically obtaining perfect focus, given access to the camera's firmware. Our method assumes the scene is Lambertian, but works reasonably well for mildly specular surfaces due to the small aperture size. While additional blur is introduced in the images due to finite sized apertures, this is only a fraction of the already existing blur due to finite pixel size. Also, use of broadband masks [18] instead of open apertures might allow make deblurring a well-posed problem. Our ray-based analysis does not take into consideration the diffraction due to reduced aperture, and this could be an important concern if the optical system is diffraction limited.

Future Directions: Despite the large and thorough existing literature on super-resolution methods, further research on mask-based methods may still supply useful contributions. To minimize diffraction effects after inserting aperture masks, further analysis may suggest alternate mask designs for demanding optical systems such as high quality microscopes. While our initial analysis addresses only planar scenes such as paintings, we could conceivably extend this to piecewise-planar surfaces. One may argue that super-resolution research has limited life given that the pixel resolution numbers in commercial systems are growing rapidly. However, the lens quality is not improving at the same rate, and their performance varies with lens settings in complex ways. Even if the lens is not diffraction limited, its PSF has a finite extent. We show that by modulating this lens PSF we are able to recover higher resolution images than otherwise possible. In addition, our analysis for exploiting

defocus on resolution enhancement maybe useful in other vision tasks such as depth from defocus and synthetic aperture imaging. Finally, mask-based methods might extend or augment video and spacetime super-resolution methods as well. Early work that assembled super-resolution images from video [10] relied on camera movement to produce interleaved sample grids. But by exploiting the movement of cameras, objects, and lighting, masks-based methods may enable resolution enhancements for all forms of video.

Conclusion: Ray-based analysis, simulations and experimental results each confirm that mask-based resolution enhancement methods can supply worthwhile improvements to image quality for stationary pictures. We avoid the challenging problems of precise mechanical movement or post-capture sub-pixel registration common in sensor-translation based methods. By modulating the lens PSF with aperture masks, relatively simple calculations can recover images that capture a greater portion of a scene's spatial frequencies. Despite our simple hand-made, ink-jet printed masks that are mounted manually on the camera lens' exterior, our method show strong visible reductions in high-frequency aliasing artifacts, and also recovers high-spatial frequency components of the image that the fixed sensor could not record in a single image.

7. Acknowledgments

We thank the anonymous reviewers for their insightful comments. This work was made possible in part by funding from the National Science Foundation (NSF) under Grants NSF-IIS-0535236 and NSF-CCF-0645973, and also by support and ongoing collaboration with Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA.

References

[1] E. H. Adelson and J. Y. A. Wang. Single lens stereo with a plenoptic camera. *Transactions on Pattern Analysis and Machine Intelligence*, 14(2):99–106, 1992.

[2] E. Chaudhuri, S. *Super-Resolution Imaging*. Kluwer Academic, 2001.

[3] D. Donoho. Compressed sensing. *IEEE Trans. on Information Theory*, pages 1289–1306, apr 2006.

[4] M. Elad and A. Feuer. Restoration of a single super-resolution image from several blurred, noisy, and undersampled measured images. *IEEE Transactions on Image Processing*, 6:1646–1658, Dec. 1997.

[5] H. Farid. *Range Estimation by Optical Differntiation*. PhD thesis, University of Pennsylvania, 1997.

[6] R. Fergus, A. Torralba, and W. T. Freeman. Random lens imaging. Technical report, MIT CSAIL, 2006.

[7] S. R. Gottesman and E. E. Fenimore. New family of binary arrays for coded aperture imaging. *Applied Optics*, 28(20):4344–4352, Oct 1989.

[8] M. Harwit and N. Sloane. *Hadamard Transform Optics*. Academic Press, 1979.

[9] M. Irani and S. Peleg. Super resolution from image sequences. *ICPR-C*, pages 115–120, 1990.

[10] M. Irani and S. Peleg. Improving resolution by image registration. *CVGIP: Graph. Models Image Process.*, 53(3):231–239, 1991.

[11] M. V. Joshi, S. Chaudhuri, and R. Panuganti. Super-resolution imaging: Use of zoom as a cue. *Image and Vision Computing*, 22(14):1185–1196, 2004.

[12] D. Keren, S. Peleg, and R. Brada. Image sequence enhancement using sub-pixel displacements. *CVPR*, pages 742–746, 1988.

[13] S. Kim, N. Bose, and H. Valenzuela. Recursive reconstruction of high resolution image from noisy undersampled multiframes. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38:1013–1027, 1990.

[14] S. Kim and W.-Y. Su. Recursive high-resolution reconstruction of blurred multiframe images. *IEEE Transactions on Image Processing*, 2:534–539, Oct. 1993.

[15] T. Komatsu, T. Igarashi, K. Aizawa, and T. Saito. Very high resolution imaging scheme with multiple different-aperture cameras. *Signal Processing: Image Communication.*, 5(5-6):511–526, 1993.

[16] S. C. Park, M. K. Park, and M. G. Kang. Super-resolution image reconstruction: a technical overview. *Signal Processing Magazine, IEEE*, 20(3):21–36, 2003.

[17] D. Rajan, S. Chaudhuri, and M. Joshi. Multi-objective super resolution: concepts and examples. *Signal Processing Magazine, IEEE*, 20(3):49–61, 2003.

[18] R. Raskar, A. Agrawal, and J. Tumblin. Coded exposure photography: motion deblurring using fluttered shutter. In *SIGGRAPH '06*, pages 795–804, 2006.

[19] G. K. Skinner. X-ray imaging with coded masks. *Scientific American*, 259:84, Aug. 1988.

[20] R. Y. Tsai and T. S. Huang. Multiframe image restoration and registration. *Advances in Computer Vision and Image Processing*, pages 317–339, 1984.

[21] P. Vandewalle, S. Süsstrunk, and M. Vetterli. A frequency domain approach to registration of aliased images with application to super-resolution. *EURASIP Journal on Applied Signal Processing*, 2006.

[22] A. Zomet and S. Nayar. Lensless imaging with a controllable aperture. In *CVPR*, pages 339–346, 2006.