Application of Structured Illumination in Nano-Scale Vision

J. Ryu[1], B.K.P. Horn[2],   [1] Harvard–MIT Div. of Health Sci. & Tech., Cambridge, MA.
M.S. Mermelstein[3],   [2] Dep. of Elec. Eng. and Comp. Sci., MIT, Cambridge, MA.
S. Hong[2], and D.M. Freeman[1,2]   [3] Lightwave Instruments, LLC, Watertown, MA.

Abstract

We describe how structured illumination patterns can be used to increase the resolution of an imaging system for optical microscopy. A target is illuminated by a sequence of finely textured light patterns formed by the interference of multiple coherent beams. The sequence of brightness values reported from a *single* pixel of a CCD imager encodes the target contrast pattern with sub-pixel resolution. Fourier domain components at spatial frequencies contained in the probing illumination patterns can be recovered from the pixel brightness sequence by solving an over-determined set of linear equations.

We show that uniform angular spacing of the beams generating the illumination patterns leads to less than ideal sampling of the transform space and we propose alternative geometric arrangements. We describe an image reconstruction algorithm based on the Voronoi diagram that applies when the transform domain is not sampled uniformly. Finally, the contrast patterns within individual pixels can be spliced together to form an image encompassing multiple pixels.

1. Introduction

Optical imaging is a relatively fast and low-cost microscopic visualization technique when compared with other imaging modalities such as electron microscopy (EM) and scanning probe microscopy (SPM). More importantly, optical imaging is non-destructive and does not require additional sample preparation or mechanical contact between a probe and the target. However, the spatial resolution of traditional optical miscroscopy is a limitation when applied to nano-scale problems. Resolution in optical imaging is fundamentally limited by the wavelength of light [1]. The smallest feature that can be resolved by optical imaging (as defined by the Rayleigh criterion [4]) is $0.61\lambda/NA$ where λ is the wavelength of light, and NA is the numerical aperture of the objective lens.

The limiting resolution predicted by this formula improves when the wavelength is reduced. However, scaling an optical imaging system to short wavelengths of light is a challenging problem. Few materials are suitable for refractive elements at very short wavelengths. Commonly used refractive materials become opaque at short wavelengths, while others exhibit birefringence.

One alternative is to use reflective elements instead of refractive elements for image formation, since some common materials reflect light of even relatively short wavelength satisfactorily. However, few examples of quality high NA imaging systems consisting entirely of reflective elements exist.

We describe an optical imaging method that combines high resolution textured illumination with a relatively low resolution sensor system, as proposed by Mermelstein [9] (Typical imaging systems instead use "low resolution" — or uniform — illumination combined with high resolution sensing). In the new system described here, a high resolution image of the target is generated computationally by processing a series of low resolution images obtained by illuminating the target with a sequence of finely textured patterns.

The resolution of such a system is limited by the resolution of the illumination pattern, and thus is still limited by the wavelength of light. However, it is not limited by the resolution of optical elements on the sensor side (such as the NA of the objective, and the sensor pixel size). More importantly, scaling to short wavelength of light becomes more practical since reflective optics can be used to generate the structured illumination.

Low resolution imaging (when a low NA objective is used) brings with it several additional advantages, such as increased working distance, wider field of view, and larger depth of field. In contrast, in a traditional system using a high NA objective, high resolution is achieved at the cost of a relatively small field of view, short working distance, and small depth of field.

A practical situation where this approach is particularly useful is that of fluorescence imaging, since in this case the wavelength of the illumination is shorter than that of the emitted light. As a result, reflective optics may be used to produce the finely textured illumination patterns, while the usual refractive optics may be used on the sensor side. Most importantly, the resolution of the new method is limited by the wavelength of the *illumination* in contrast to the usual fluorescence imaging method
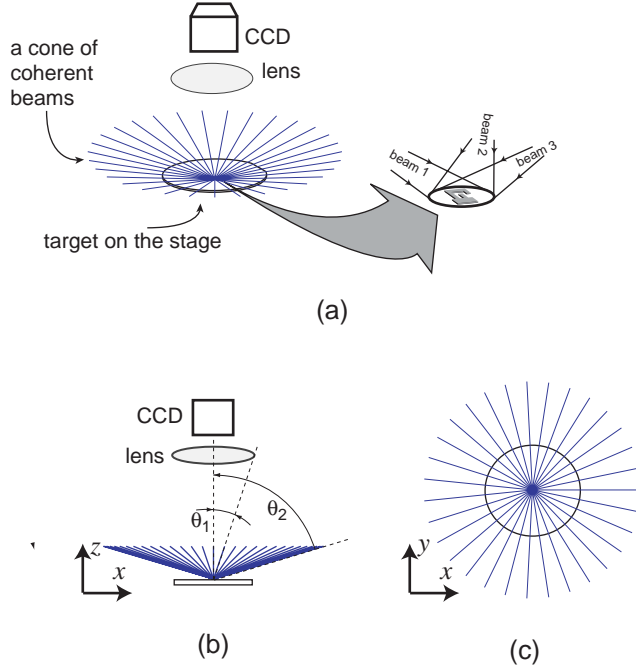
1

Figure 1: Generation of structured illumination using the interference of multiple coherent beams. (a) Schematic of the imaging setup showing a lens, a CCD, and a cone of 31 coherent beams illuminating the target. The target is placed where the beams overlap, as illustrated in the right side. For clarity, only three beams are shown. (b) Side view of the setup. The cone half angle of the beams (θ_2) and the cone half angle of the lens (θ_1) are indicated. (c) A view of the cone of beams looking from the top.

where resolution is limited by the (longer) wavelength of the *emitted* radiation.

Structured illumination has been successfully used in range and shape imaging [3] [8] and photographic scene analysis [6]. Also, there have been recent successes in using structured illumination to improve the resolution of fluorescence microscopy [7] [5]. In that work, however, the structured illumination is a simple periodic pattern formed by the interference of two laser beams (or two pairs of opposite beams). We instead use complex texture patterns created by the interference of *many* laser beams as practiced by Mermelstein [9]. A series of different textured patterns are produced by varying the relative phases and amplitudes of the set of beams.

In this paper, we present the framework of the method and associated computational schemes. Physical apparatus to generate a controlled sequence of finely textured structured illumination patterns will be described in a separate paper. First we show how a sequence of structured illumination patterns can be used to encode target contrast information in a sequence of brightness values reported from a single pixel of a charge-coupled device (CCD). Next, computational methods to decode this

information, and to reconstruct an image of the target within a small area are presented, together with a method for combining images from multiple pixels. Finally, we explore a method to enhance the quality of the reconstructed image by modifying the angular placement of the laser beams.

2. High Resolution Structured Illumination

High spatial resolution structured illumination patterns are formed by the interference of multiple coherent beams. Fig. 1a illustrates an optical imaging setup using a lens and a CCD in which illumination is provided by a number of coherent beams that overlap in an area where a target is placed. In the region of overlap, a light pattern is formed as the result of interference between the beams. It is this pattern that illuminates the target. The right side of Fig. 1a shows a magnified view of the target region where beams overlap. Only three beams are shown for clarity. The overlapping beams form a cone in three-dimensional space. Fig. 1b is a side view of the cone of beams, while Fig. 1c shows a view from the top (i.e. from the direction of the lens and the CCD imager).

Suppose there are N coherent beams overlapping in the target region. The electric field, $\mathbf{E}$, of the n-th beam as a function of position $\mathbf{r}$ and time t can be written as

$$\mathbf{E}_n(\mathbf{r},t) = a_n \cos(\mathbf{k}_n \cdot \mathbf{r} - \omega_n t + \phi_n)\hat{\mathbf{p}}_n \qquad (1)$$

where a_n is the amplitude, $\mathbf{k}_n$ the wavenumber[1], ω_n the optical frequency, ϕ_n the optical phase, and $\hat{\mathbf{p}}_n$ the polarization vector. The intensity pattern formed by the interference of N beams then is

$$I(\mathbf{r},t) = \left| \sum_{n=1}^{N} \mathbf{E}_n(\mathbf{r},t) \right|^2 . \qquad (2)$$

In the method presented here, the beams all ultimately originate from the same source and so λ_n and ω_n are the same for all beams. In addition, the polarization vectors $\hat{\mathbf{p}}_n$ of the beams are assumed to be equal[2]. Then the time average of $I(\mathbf{r},t)$ over one cycle of the electromagnetic oscillation can be written

$$I(\mathbf{r}) = \frac{1}{2} \sum_{l=1}^{N} \sum_{m=1}^{N} a_l a_m \cos(\mathbf{k}_{lm} \cdot \mathbf{r} + \phi_{lm}) \qquad (3)$$

---

[1]The wavenumber $\mathbf{k}_n$ is $(2\pi/\lambda_n)\hat{\mathbf{l}}_n$, where λ_n is the wavelength and $\hat{\mathbf{l}}_n$ is a unit vector defining the direction of propagation of the n-th beam. The vectors $\mathbf{k}_n$ and $\hat{\mathbf{p}}_n$ are perpendicular to one another.

[2]In fact, the polarization vectors of the beams cannot be made exactly equal, but this is a good approximation if the beams are TM (Transverse Magnetic) polarized and the cone half angle of the beams (θ_2) is close to 90°.
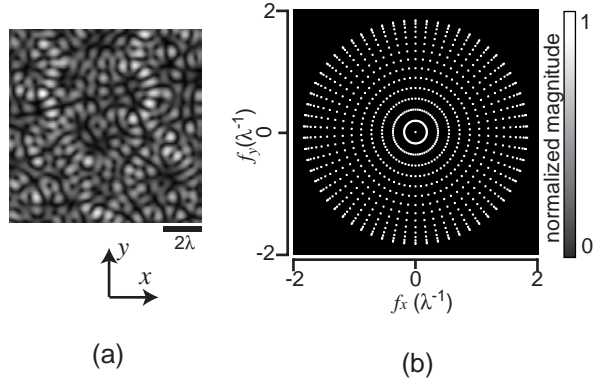
Figure 2: (a) Computed structured illumination pattern in the target plane formed by the cone of 31 beams shown in figure 1 ($\theta_2 = 78^\circ$). (b) Fourier transform magnitude of the pattern in (a) showing a total of 931 bright spots corresponding to a set of spatial frequency components contained in the intensity pattern in (a).

where $\mathbf{k}_{lm} = \mathbf{k}_l - \mathbf{k}_m$ and $\phi_{lm} = \phi_l - \phi_m$. Eq. (3) shows that the intensity pattern is a function of position $\mathbf{r}$ and is completely characterized by the amplitudes (a_n), the optical phases (ϕ_n), and the wavenumbers ($\mathbf{k}_n$) of the beams. When the directions of propagation of the beams are fixed in space by the physical arrangement of the apparatus, then the intensity pattern $I(\mathbf{r})$ becomes a function of just the amplitudes and optical phases of the beams. Consequently, the intensity pattern can be changed by modulating these beam parameters. Fig. 2a shows a computer simulated intensity pattern in the target plane formed by the interference of 31 beams as illustrated in Fig. 1. The cone half angle of the beams (θ_2) was chosen to be 78° in this illustration. In this particular example, the amplitudes of all beams were set equal and the optical phases of the beams were chosen randomly between 0 and 2π with uniform distribution.

Eq. (3) shows that in the Fourier domain, the intensity pattern $I(\mathbf{r})$ is composed of a set of distinct spatial frequency components $\mathbf{k}_{lm}$, for $l = 1, 2, \cdots, N$ and $m = 1, 2, \cdots, N$. The total number of distinct frequency components (p) is determined by the number of beams (N), and is $p = N^2 - N + 1$.[3] The values of the spatial frequencies themselves are determined by the differences between the wavenumbers of all possible pairs of beams. The highest spatial frequency is $2/(\lambda \sin(\theta_2))$. Fig. 2b shows the Fourier transform magnitude of the intensity pattern in Fig. 2a. It shows a total of 931 (i.e. $31^2 - 31 + 1$) bright spots corresponding to the spatial frequency components of the pattern in Fig. 2a.

As will be shown in the following section, the resolu-

tion of the proposed system is determined by the spatial frequency components contained in the structured illumination. Since the highest spatial frequency component (and hence the smallest feature size) of the illumination pattern is controlled by the cone half angle of the beams (θ_2) at a given wavelength of light, the angle θ_2 controls the resolution of the system. For a traditional imaging system in Fig. 1 with uniform illumination, the resolution is instead controlled by the cone half angle of the objective lens (the angle θ_1 in Fig. 1b, NA = $\sin(\theta_1)$). The ratio $\eta = \sin(\theta_2)/\sin(\theta_1)$ represents the increase in maximum spatial frequency content present as the result of the use of structured illumination.

In computer simulations that follow, the NA of the lens in Fig. 1 was chosen to be 0.13 (therefore, $\theta_1 = 7.5^\circ$), representing the actual NA of a lens from a leading manufacturer of microscope objectives. This lens selection aims at advantages in field of view, working distance, and depth of field that a low NA lens offers. The cone half angle of the beams (θ_2) was selected to be 78°, representing the corresponding value in the multiple beam interference pattern projector we have designed and built. Notice that the value η — representing the increase in maximum spatial frequency content present due to structured illumination — corresponds here to $\sin(78^\circ)/\sin(7.5^\circ) = 7.5$.

Note that the three-dimensional brightness pattern $I(\mathbf{r})$ is constant in the z-direction if all the wavenumbers $\mathbf{k}_n$ have the same z-component. This happens when the beams all come in at the same angle ($\pi/2 - \theta_2$ in Fig. 1) with respect to the target plane $z = 0$. The result is very large depth of field (limited only by the low NA of the objective lens). As a result we treat the target contrast as an essentially two-dimensional pattern in the following (i.e. a function of $\mathbf{r}$ in the target plane $z = 0$).

3. Encoding and Decoding Target Information

Suppose that there are N beams and M distinct structured illumination patterns are projected in sequence by controlling the amplitudes (a_n) and the optical phases (ϕ_n) of the beams. Let $I_j(\mathbf{r})$ represent the intensity of the j-th illumination pattern as a function of position $\mathbf{r}$. The planar target is represented by a contrast value as a function of position, $C(\mathbf{r})$.[4] In a simulation we use the image shown in Fig. 3a as the target contrast pattern. The light pattern that falls on the CCD imager, $g_j(\mathbf{r})$, is given by

$$g_j(\mathbf{r}) = [C(\mathbf{r})I_j(\mathbf{r})] * h(\mathbf{r}) \quad j = 1, 2, \cdots, M \quad (4)$$

---

[3]When $l = m$ in Eq. (3), the corresponding spatial frequency $\mathbf{k}_{lm}$ becomes zero (DC), so N of the N^2 combinations contribute to a single frequency component.

[4]Here, the position vector $\mathbf{r}$ is in target coordinates. For convenience, the magnification of the imager can be treated as if it were unity, in which case the position vector in the target coordinates also designates the position in the image coordinates.

where $h(\mathbf{r})$ represents the point spread function (PSF) of the lens [4] and $*$ designates the convolution operation.

The CCD can be viewed as an array of light sensors where each pixel corresponds to a single light sensor. The brightness b_j reported from a single pixel of the CCD imager positioned at the origin ($\mathbf{r} = 0$) of the coordinate system becomes

$$b_j = \int g_j(\mathbf{r}) p(\mathbf{r}) \, d\mathbf{r} \quad j = 1, 2, \cdots, M \quad (5)$$

where $p(\mathbf{r})$ represents the response of the pixel, modelled as a two dimensional rectangular function which is one inside the area of the pixel and zero elsewhere. Substituting Eq. (4) into Eq. (5) and preforming some algebraic manipulations leads to

$$\begin{aligned} b_j &= \int [w(\mathbf{r}) C(\mathbf{r})] I_j(\mathbf{r}) \, d\mathbf{r} & (6) \\ &= \int C_w(\mathbf{r}) I_j(\mathbf{r}) \, d\mathbf{r} \quad j = 1, 2, \cdots, M & (7) \end{aligned}$$

where $w(\mathbf{r}) = h(\mathbf{r}) * p(\mathbf{r})$ is the two-dimensional convolution of $h(\mathbf{r})$ and $p(\mathbf{r})$ (i.e. the combined PSF of the lens and the CCD sensor element). In the following, $w(\mathbf{r})$ will be called the "window" function — it gives the reponse of a picture cell to incident light as a function of position relative to the center of that picture cell. A simulated windowing function $w(\mathbf{r})$ is shown in Fig. 3b. In Eq. 7, $C_w(\mathbf{r}) = w(\mathbf{r}) C(\mathbf{r})$ is the "windowed" target contrast function. A sample $C_w(\mathbf{r})$ shown in Fig. 3c corresponds to $C(\mathbf{r})$ and $w(\mathbf{r})$ shown in parts (a) and (b).

Eq. (7) shows that, for a given sequence of textured illumination patterns, the recorded sequence of brightness values is a function of $C_w(\mathbf{r})$. For a computer-simulated sequence of 1400 illumination patterns[5], and the target shown in Fig. 3a, the resulting brightness sequence is shown in Fig. 3d. The sequence of illumination patterns was generated by changing the optical phases of the beams.[6]

The sequence of brightness values $b_j, j = 1, 2, \cdots, M$, in Eq. (7) encodes information on the windowed target contrast $C_w(\mathbf{r})$ at spatial frequencies present in the sequence of illumination patterns $I_j(\mathbf{r}), j = 1, 2, \cdots, M$. This can be more clearly seen after replacing $I_j(\mathbf{r})$ in Eq. (7) by the expression in Eq. (3) and performing several steps of algebraic manipulations to rewrite Eq. (7) as the following matrix equation [9]:

$$\mathbf{b} = A\mathbf{x} \quad (8)$$

[5]We picked $M = 1400$ patterns so as to yield an overdetermined system of linear equations in the unknown components at 931 frequencies.

[6]For each illumination pattern, 31 optical phases were chosen randomly between 0 and 2π with uniform distribution. In this simulation the amplitudes of the 31 beams were kept equal.
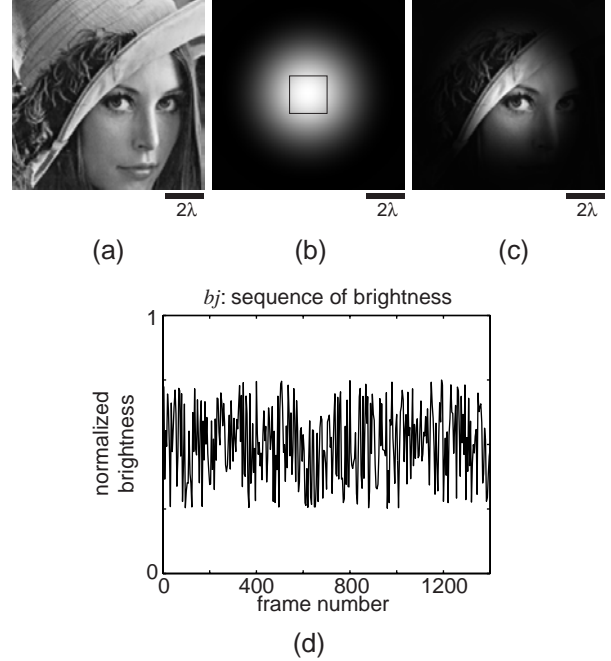

2λ   2λ   2λ

(a)   (b)   (c)

b_j: sequence of brightness

(d)

Figure 3: Encoding sub-pixel target contrast information in a sequence of brightness values from a single pixel. (a) $C(\mathbf{r})$; a target contrast function. (b) $w(\mathbf{r})$; a window function for a pixel positioned at the origin (the center of the image). NA = 0.13, pixel size = 2λ. The rectangle at the center shows the actual size of the pixel relative to the window function. (c) $C_w(\mathbf{r}) = C(\mathbf{r}) w(\mathbf{r})$; a windowed target, (d) Simulated sequence of brightness values, $\mathbf{b}_j$, reported from a single pixel while the target was illuminated by a sequence of $M = 1400$ different structured illumination patterns.

where the $M \times 1$ column vector $\mathbf{b} = [b_1, b_2, \cdots, b_M]^T$ is the measured brightness sequence, while the $p \times 1$ column vector $\mathbf{x}$ equals

$$[x_{DC}, x_{12}^c, x_{12}^s, x_{13}^c, x_{13}^s, \cdots, x_{(N-1)N}^c, x_{(N-1)N}^s]^T$$

and represents a total of $p = (N^2 - N + 1)$ Fourier transform coefficients of the windowed target contrast $C_w(\mathbf{r})$ at spatial frequencies present in the illumination pattern. Specifically,

$$x_{DC} = (N/2) \int C_w(\mathbf{r}) \, d\mathbf{r} \quad (9)$$

$$x_{lm}^c = \int C_w(\mathbf{r}) \cos(\mathbf{k}_{lm} \cdot \mathbf{r}) \, d\mathbf{r} \quad (10)$$

$$x_{lm}^s = \int C_w(\mathbf{r}) \sin(\mathbf{k}_{lm} \cdot \mathbf{r}) \, d\mathbf{r} \quad (11)$$

The j-th row, A_j, of the $M \times p$ matrix A, is determined by the amplitudes (a_n) and the optical phases (ϕ_n) of the beams for the j-th illumination. Specifically, A_j equals

$$[1, \alpha_{12}, -\beta_{12}, \alpha_{13}, -\beta_{13}, \cdots, \alpha_{(N-1)N}, -\beta_{(N-1)N}]$$

4

$f_y(\lambda^{-1})$

$f_x(\lambda^{-1})$
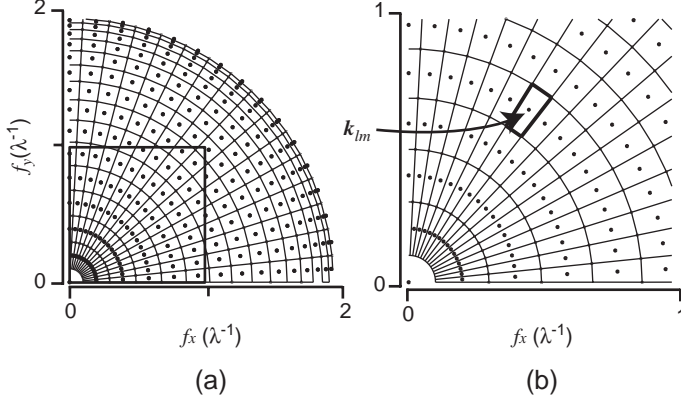
(a)

k_{lm}

$f_x(\lambda^{-1})$

(b)

Figure 4: Calculating the weights to be given to samples in Fourier space using the areas of the Voronoi polygons. (a) The Voronoi diagram for the distribution of the Fourier sample points in Fig. 2b. Only the first quadrant of the Fourier space is shown for clarity (The point at the lower left corner corresponds to DC). (b) Magnified view of a rectangular sub-region in (a) indicating an example Fourier sample point ($\mathbf{k}_{lm}$) and its corresponding Voronoi polygon.

with $\alpha_{lm} = a_l a_m \cos(\phi_{lm})$ and $\beta_{lm} = a_l a_m \sin(\phi_{lm})$.

For $M > p$, Eq. (8) is an overdetermined linear system. Using knowledge of the illumination (i.e. the matrix A) and the measured brightness sequence (the vector $\mathbf{b}$), we can find $\overline{\mathbf{x}}$, the best approximation to the solution $\mathbf{x}$ of Eq. (8) in the least squares sense, using

$$\overline{\mathbf{x}} = (A^T A)^{-1} A^T \mathbf{b} \tag{12}$$

where $(A^T A)^{-1} A^T$ is the pseudoinverse of A.

Notice that the number of estimated Fourier transform coefficients (p) increases quadratically with the number of beams (N). This means that when one uses a relatively modest number of beams (e.g. 31), a large enough number of Fourier transform coefficients (e.g. 931) of the target contrast pattern can be estimated to enable the reconstruction of an image of the target from the data[7].

Notice also that the resolution (or equivalently the spatial frequency content) of such an image is determined by the illumination and is not directly limited by the factors limiting resolution in a traditional imaging system, such as the NA of the lens, and the CCD pixel size. Instead of determining the resolution, the NA of the lens controls the size of the window function $w(\mathbf{r})$ and therefore defines the area of the target that is represented by the vector $\mathbf{x}$ in Eq. (8).

---

[7]As a rough rule of thumb, N beams provide about enough information to reconstruct the N^2 brightness values in an $N \times N$ array of sub-pixel cells.

4. Image Reconstruction

If sufficiently many Fourier transform coefficients are obtained using the method described in the previous section, then an image of the windowed target can be reconstructed. As is well known, an arbitrary function can be reconstructed exactly from its Fourier transform. The Fourier transform is a continuous function of frequency. In contrast, here only a finite number of samples of the Fourier transform of $C_w(\mathbf{r})$ are obtained by solving Eq. (8). Furthermore, as is illustrated in Fig. 2b, the distribution of these 'Fourier sample points' is highly non-uniform. First of all, the sampling density varies with frequency, being low at mid-frequencies and high at both low and high frequencies. Secondly, in the low spatial frequency region, while the sample points are densely populated along the angular direction, they are sparsely populated along the radial direction. The reverse is true at high frequencies.

To compensate for sampling density non-uniformity in Fourier space, each sample point in the Fourier space is weighted inversely with the density of sample points in its vicinity for purposes of image reconstruction. This notion was implemented by first allocating non-overlapping space-filling elemental areas to the sample-points. The density at a sample point is inversely proportional to the corresponding elemental area. These areas in turn are defined by calculating the Voronoi polygon for each sample point. The Voronoi polygon is defined by drawing a boundary enclosing all points lying closer to the sample point in question than to any other sample point [2]. Fig. 4a shows the Voronoi diagram of the sample points in Fig. 2b, which is the set of all Voronoi polygons for all sample points. For clarity, only the first quadrant of the Fourier space is shown. Fig. 4b is an enlarged view of the sub-region of (a) indicated by a rectangle. Sample points are shown as dots, and borderlines between adjacent Voronoi polygons are marked as lines. Notice that the polygons are more elongated in regions of low as well as at high frequencies.

The area of each Voronoi polygon was calculated and used as a weighting factor for the corresponding sample point. An image of the estimated windowed target $\widehat{C}^w(\mathbf{r})$ was then reconstructed by summing all weighted frequency components as in Eq. (13).

$$\widehat{C}^w(\mathbf{r}) = s_{DC} x_0 + 2 \sum_{l=1}^{N-1} \sum_{m=l+1}^{N} s(\mathbf{k}_{lm}) \chi_{lm} \tag{13}$$

where s_{DC} and $s(\mathbf{k}_{lm})$ are the areas of the Voronoi polygon corresponding to DC and the spatial frequency $\mathbf{k}_{lm}$, respectively and

$$\chi_{lm} = x_{lm}^c \cos(\mathbf{k}_{lm} \cdot \mathbf{r}) + x_{lm}^s \sin(\mathbf{k}_{lm} \cdot \mathbf{r}). \tag{14}$$
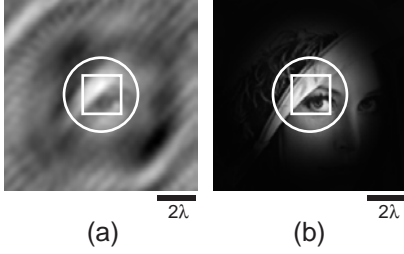
5

Figure 5: (a) An image of the windowed target $C_w(\mathbf{r})$ reconstructed by processing the brightness sequence in Fig. 3d by the method described in this paper. (b) The original windowed target shown as a reference.

The above reconstruction formula treats the measured energy contributions as being localized at a single frequency but considers them as representative of the energy over the corresponding Voronoi polygon.[8]

Fig. 5a shows an image reconstructed by processing the pixel brightness sequence in Fig. 3d as described above. The area in the image itself that looks approximately circular defines a region where the window function in Fig. 3b is more than 50% of its maximum value. The rectangular region inside the approximately circular area is the pixel whose brightness measurements were used in reconstructing this image. Part (b) is the original windowed target shown as a reference. Notice that the reconstructed image within the more or less circular area in part (a) is similar to the corresponding region in part (b) — with some loss of sharpness. Fig. 5 clearly demonstrates that, using a sequence of brightness measurements from a single pixel, an image can be reconstructed with sub-pixel detail within that pixel. In fact, as shown, an area somewhat larger than the pixel can be reconstructed.

Two factors control the quality of the reconstructed image. First, the wavelength λ of light and the cone half angle of beams θ_2 determines the maximum spatial frequency $(2\pi/(\lambda \sin(\theta_2)))$ present in the illumination and therefore controls the sharpness of the reconstructed image. Second, the total number of beams (N) determines the number of Fourier sample points and therefore controls the density of sampling in the Fourier domain. This in turn determines how many substantially independent sub-pixel samples can reasonably be expected to be recovered.

---

[8]We have measurements only at discrete unequally spaced frequencies in the transform domain. We could assume as an approximation that the transform was constant in each Voronoi polygon. Instead we consider all the energy for that polygon to be concentrated at a point.
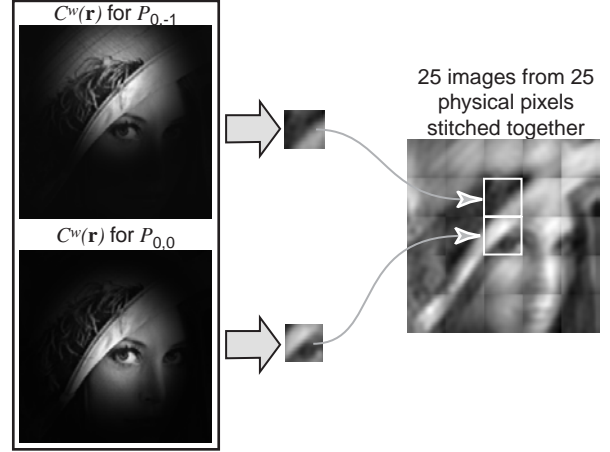

Figure 6: Combining images from multiple pixels (The dropoff in brightness at the edges of the tiles is due to dropoff in the windowing function — which not been compensated for).

5. Combining Multiple Pixels

In the previous section, it was shown that a (small) image can be generated using data from a *single* pixel of a CCD. When an array of pixels of a CCD are considered, the situation is equivalent to multiple imagers working in parallel — each looking at a slightly different area. Images generated from multiple pixels can be combined to form a larger image.

Eq. (7) describes the brightness of a pixel positioned at the origin of the coordinate system in terms of the window function, $w(\mathbf{r})$, the target contrast, $C(\mathbf{r})$, and the illumination, $I_j(\mathbf{r})$. For a pixel whose center is positioned at $\mathbf{r} = \mathbf{r}_c$, the brightness of the pixel is now described by

$$b_j = \int [w(\mathbf{r} - \mathbf{r}_c)C(\mathbf{r})]I_j(\mathbf{r})\,d\mathbf{r} \qquad (15)$$

where $w(\mathbf{r} - \mathbf{r}_c)$ is $w(\mathbf{r})$ shifted in space by $\mathbf{r}_c$. The windowing function shifted by the vector $\mathbf{r}_c$ defines a sub-region in the target space to which the output of a particular pixel in the CCD is responsive (and conversely, the area in which a contrast function can be reconstructed from the coded information obtained from the pixel).

In this paper, a simple method for combining images from multiple pixels is presented. Fig. 6 illustrates this procedure using two adjacent pixels, $P_{0,-1}$ and $P_{0,0}$. The two images inside the left box of Fig. 6 represent the target contrast pattern multiplied by two different window functions corresponding to the two chosen pixels. Processing the data from each pixel is equivalent to performing Fourier analysis of the corresponding windowed target. The figure illustrates how the position of the window corresponds to the positions of the pixels.

When processed, the data from each pixel generates an image of a windowed portion of the target contrast. From
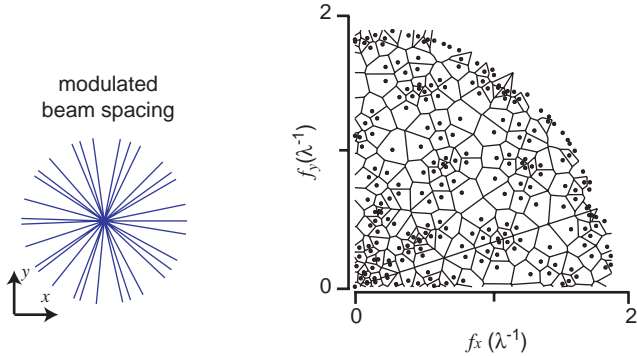
Figure 7: A different placement of 31 beams from that in Fig. 1c — with sinusoidally modulated spacings between adjacent beams (left) — and the resulting distribution of sample points in Fourier space along with the corresponding Voronoi diagram (right).

Figure 8: Images of the target reconstructed from data obtained (a) using uniform beam spacing, and (b) using modulated beam spacing.

each such "sub-image," a region corresponding to the area of the pixel is cropped (for example, the rectangular area defined in Fig. 5a). The two small images next to the left box in Fig. 6 are sample cropped images. The result of stitching the two small images together is shown on the right. That image is a mosaic of a total of 25 small images generated from 25 individual pixels.

6. Optimizing Beam Placement

The distribution of Fourier sample points is determined by the angular arrangement of the beams. The distribution in Fig. 2b results from the beam placement shown in Fig. 1c with uniform angular spacing of beams. In this case, the sample points in Fourier space lie on the intersections of a group of concentric circles and a group of radial lines. The corresponding Voronoi diagram (Fig. 4) shows that the Voronoi polygons in the low and high frequency regions have very thin shapes (i.e. high aspect ratio), which is an indication of sampling density non-uniformity in those regions. The reason this is undesirable is that for fixed cell area, an elongated cell combines frequency components that are further apart than does a cell of a more "rounded" shape.

One can try and find a set of beam angles that provides a more uniform distribution of sample points in Fourier space. One can for example minimize some measure of the elongation of the Voronoi polygons using numerical optimization. Fig. 7 illustrates a particular example of non-uniform beam spacing. The left part of Fig. 7 shows beam placement (with the same 31 beams) in which the angles between adjacent beams vary sinusoidally from beam to beam — rather than all being the same. The right part shows the resulting distribution of sample points in Fourier space and the corresponding Voronoi diagram — shown here in the first quadrant only. Compared to
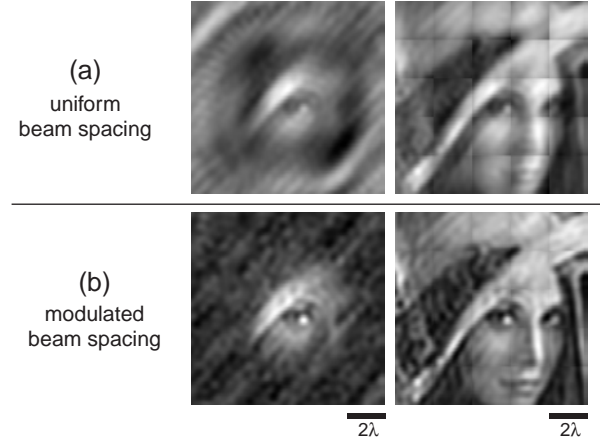
the uniform beam spacing case illustrated in Fig. 4a, the resulting Voronoi polygons — while still showing considerable variation in area — exhibit significantly reduced aspect ratios, especially in the low and high spatial frequency regions, indicating better sampling.

Fig. 8 compares images reconstructed using data for the two different beam placements considered here. Part (a) shows two images for the uniform beam spacing case. The left side is an image reconstructed using data from a single pixel located at the center, while the right side is the result of combining images from 25 pixels. Part (b) shows corresponding images for the modulated beam spacing case. The effect of improving the distribution of sample points in Fourier space is clear from the reconstructed images. Modulated beam spacing produces images with lower energy background artifacts and improved sharpness.

Finally, images of the same target acquired with different conditions are compared in Fig. 9. The image in part (a) simulates an image acquired with a 0.13 NA lens, using uniform illumination. Here the size of the pixel of the CCD imager was assumed to be infinitesimally small. Therefore, part (a) corresponds to the theoretical resolution limit of the 0.13 NA lens. The image in part (b) simulates the same condition as in part (a), but with finite pixel size of 2λ. The two images in part (c) and part (d) are the same as part (a) and part (b), but with a 0.98 NA lens that has the same cone half angle as the cone half angle of the beams (78º). The image in part (e) was generated with a 0.13 NA lens and a finite pixel size (2λ), but using a structured illumination sequence.

Several observations can be made here: First, the results in Fig. 9 clearly illustrate that the structured illumination technique overcomes the practical resolution limit
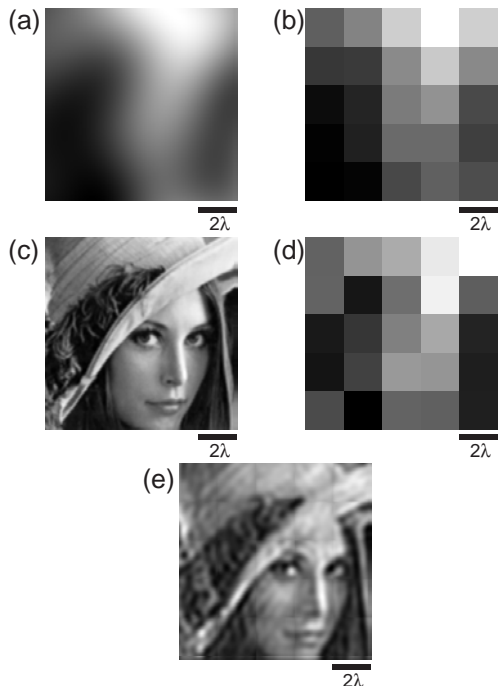
Figure 9: Images of the same target acquired with different conditions. See text for details.

set by the finite pixel size of the CCD imager, generating an image of the target with sub-pixel details. Second, the resolution of the technique is comparable to the theoretical resolution limit of a 0.98 NA lens — i.e. a lens with the same cone half angle as the cone half angle of the coherent beams — although still limited by the wavelength of light. An important factor, however, is that, unlike the 0.98 NA objective lens of the sensor, the illumination part of the new system can be made entirely of reflective optical elements, which makes scaling to short wavelength of light more practical. Finally, the method has advantages in terms of the field of view, the working distance, and the depth of field compared with a 0.98 NA lens, since these parameter are here instead determined by the 0.13 NA lens on the sensor side.

7. Conclusion

We described an optical microscopy method in which high resolution illumination in the form of a sequence of finely textured light patterns is combined with low resolution sensing. We showed that the sequence of brightness values reported from a single pixel encodes the target contrast pattern within (and near) that pixel. A computational scheme to process non-uniformly spaced samples in Fourier space to produce a reconstructed image of the target contrast within (and near) the pixel was presented. The resolution of such a system is primarily limited by the

illumination system, not by the NA of the objective or the size of the pixels, as in traditional microscopic imaging methods. Since the physical apparatus to generate the illumination can be made entirely of reflective optical elements, the method holds potential for scaling optical microscopy to very short wavelengths of light.

In future work, we plan to find "optimal" arrangements of beam angles using numerical optimization technique. To do this, we will have to develop suitable optimization criteria. Work on antenna placement in radio astronomy may be relevant in this connection [10]. We also plan to investigate improved methods for combining images from multiple pixels, especially noting that windows for adjacent pixels overlap one another significantly. The method described here is similar in concept to the short-time Fourier transform (STFT) analysis and synthesis technique in that Fourier analysis is performed using overlapping windows that are shifted versions of one another. Formulating the method as a variation of STFT may provide new insights.

References

[1] E. Abbe, *Arch. Mikrosk. Anat.*, vol. 9, pp. 413–468, 1873.

[2] F. Aurenhammer. "Voronoi Diagrams: A Survey of a Fundamental Geometric Data Structure." *ACM Computing Surveys*, 23:345-405, 1991.

[3] Paul Besl. *Advances in Machine Vision*, chapter 1 – Active optical range imaging sensors, pp. 1–63. Springer-Verlag, 1989.

[4] M. Born, E. Wolf, *Principles of Optics*, Cambridge University Press, New York, 1999.

[5] J.T. Frohn, H.F. Knapp, A. Stemmer, "True optical resolution beyond the Rayleigh limit achieved by standing wave illumination", *Proc. Natl. Acad. Sci. USA*, vol. 97, pp. 7232–7236, 2000.

[6] A.E. Grace, D. Pycock, H.T. Tillotson, M.S. Snaith, "Active shape from stereo for highway inspection", *Machine Vision and Applications*, vol. 12, no. 1, pp. 7–15, 2000.

[7] M.G.L. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy", *J. Microsc.*, vol. 198, pp. 82–87, 2000.

[8] R.A. Jarvis, "A perspective on range-finding techniques for computer vision", *IEEE Trans. Pattern Analysis Mach. Intell.*, vol. 5, pp. 122–139, 1983.

[9] M.S. Mermelstein, PhD thesis, Massachusetts Institute of Technology, 2000.

[10] E. Keto, "The Shapes of Cross-Correlation Interferometers," *The Astrophysical Journal*, Vol. 475, pp. 843–852, 1997.

8