

# Super-Resolution Image Synthesis using Projections onto Convex Sets in the Frequency Domain<sup>†</sup>

Frederick W. Wheeler<sup>\*a</sup>, Ralph T. Hoctor<sup>a</sup> and Eamon B. Barrett<sup>b</sup>

<sup>a</sup>General Electric Global Research, 1 Research Circle, Niskayuna, NY, USA 12309;

<sup>b</sup>Lockheed Martin Space Systems, ATC, 1111 Lockheed Martin Way, Sunnyvale, CA, USA 94089

## ABSTRACT

In this report we propose a frequency domain POCS algorithm for the canonical problem of super-resolution (SR) image synthesis. Unlike previous frequency domain SR algorithms, this approach is structured to accommodate rotations of the source relative to the imaging device, which we believe to help in producing a well-conditioned image synthesis problem. Generally, frequency domain methods have been used when component images were related by subpixel shifts only, because rotations of a sampled image do not correspond to a simple operation in the frequency domain.

**Keywords:** superresolution, projections onto convex sets, image rotation

## 1. INTRODUCTION

Images obtained from various optical imaging systems are often limited in resolution, not by the imaging optics, but by the light intensity sensors on the image formation plane. This comes about because system designers prefer larger sensors that yield higher signal-to-noise ratios for fixed time apertures. When sensors are the limiting factor in the spatial frequency response of an imaging system, there is also an associated aliasing phenomenon created. In effect, the locations of the samples from such an imaging instrument are separated by the dimension of the sensor. When this dimension is much smaller than the optical spot size of the system, spatial frequencies are limited by the optical transfer function and the image is well-sampled. When the dimension of the sensor is larger than the optical spot size, the effect is that of smoothing the image with a rectangular convolving kernel with only one sample at each non-overlapping kernel position, which produces aliasing.

In some imaging systems where the image plane sensors dominate the system impulse response, there is the possibility of collecting multiple images of the same scene. This is the case, for example, in remote sensing. In general, these images differ from one another by shifts or rotations or both. Such data sets can be processed to produce a single, high-resolution image. The process of reconstructing a high-resolution image from multiple low-resolution images is referred to as “super-resolution image reconstruction.” This is an active field of research, and its aims and methods are reviewed in two recent retrospectives<sup>1,2</sup>.

The term “super-resolution” has also been used to refer to an earlier scheme for enhanced-resolution imaging of continuous sources, that of bandlimited extrapolation of an aperture plane measurement to produce the effect of a larger aperture<sup>3</sup>. (This is equivalent to bandlimited extrapolation of the spatial frequencies of an image<sup>4</sup>.) It was proposed that this extrapolation could be performed either using a prolate spheroidal function expansion<sup>5</sup> or the Gerchberg-Papoulis algorithm<sup>6</sup>. Both computational approaches suffer from the need for very high dynamic range measurements made with very high SNR, and are not practical for more than a small increase in resolution<sup>5,7,8</sup>. (However, note that the use of *a priori* information about the source can improve this situation to some extent<sup>9,10</sup>.) This earlier “super-resolution” is not necessarily related to the present topic, and this has been a cause of some confusion. However, since the term “super-resolution” (SR) is so well established, we adopt it here.

The stated objectives of SR image synthesis are not uniform across the community working in this area and they have evolved with time. For example, the review article of Borman and Stevenson<sup>2</sup> specifically includes the aim of classical super-resolution<sup>4</sup>: extrapolation of the frequency content beyond the limits imposed by the optics. In this work

---

<sup>†</sup> Lockheed Martin Patent Pending

\* wheeler@research.ge.com

we do *not* adopt bandlimited extrapolation of spatial frequencies as a goal. To make our presentation specific, we will adopt the imaging scenario and objectives of Stark and Oskoui<sup>11</sup>: multiple images are formed and recorded for which the image PSF is much smaller than the spatial impulse response of the sensor. Rotations between the instrument and the source are allowed. The goal of the computation is to produce an image with a PSF of the same size as that of the optical system.

The observable intensity distribution on the image formation plane includes the effects of the light intensity sensors; this differs from the actual intensity distribution on the image formation plane, which includes only the effect of the optical point spread function. The observable distribution corresponding to the  $m^{\text{th}}$  image is modeled as:

$$G_m(x, y) = I(x, y) ** f(\cos \theta_m x - \sin \theta_m y + \Delta x_m, \sin \theta_m x + \cos \theta_m y + \Delta y_m) \quad (1)$$

where  $I(x, y)$  is the spatial impulse response of the imaging system,  $f(x, y)$  is the continuous source to be imaged,  $\theta_m$  is the rotation angle of the  $m^{\text{th}}$  image with respect to the coordinate system of the imaging device,  $(\Delta x_m, \Delta y_m)$  gives the shift for the  $m^{\text{th}}$  image and the double asterisk represents 2-dimensional convolution. The system impulse response is itself the convolution to two distinct components: the point spread function of the imaging optics,  $P(x, y)$ , and the impulse response of the image intensity sensor,  $S(x, y)$ :

$$I(x, y) = P(x, y) ** S(x, y) \quad (2)$$

The  $m^{\text{th}}$  observation is a sampled version of  $G_m$ , observed in additive noise:

$$X_m(k_1, k_2) = G_m(k_1 T, k_2 T) + n(k_1, k_2), \quad \text{for } k_1, k_2 = 0, \dots, (K-1) \quad (3)$$

where  $T$  is the spatial sampling interval in both the  $x$  and  $y$  directions, which is determined by the size of the imaging sensor. We will assume the images are square and sampled at the same rate in both dimensions, though the method is easily generalized to rectangular images and different sample rates in each dimension. The computation assumes knowledge of the rotation angles  $\theta_m$  and shifts  $(\Delta x_m, \Delta y_m)$ , which may be determined via image registration. The synthesized image will be sampled at a rate that exceeds the Nyquist rate for its spatial frequency content. The image formation plane intensity distribution to be synthesized is, when the instrument is aligned with the high-resolution coordinate system:

$$Y(x, y) = P_s(x, y) ** f(x, y) \quad (4)$$

and the image is sampled at the interval  $T/W$ , with  $W$  selected so that the interval is smaller than the Nyquist sample period for  $Y(x, y)$ . The symbol  $P_s(x, y)$  in (4) represents the point spread function to be synthesized, and, in general, it could differ from the PSF of the imaging system used in (2). For example, the PSF to be synthesized might be that corresponding to a uniform weighting of spatial frequencies in the image; for a circular aperture, this is the surface of rotation of  $J_1(r)/r$ . Note that this is not a PSF that can be obtained by imaging with an optical intensity imaging system; however, this synthesis could be implemented by incorporating a re-weighting of the spatial frequencies of the image in the projection operator, as described later. Of course, the actual PSF can be used as the desired PSF,  $P_s$ , in which case (4) represents the actual intensity distribution on the image formation plane. This can be implemented by using only the frequency response function associated with the image sensor in the projection operator.

It is important to note the differences between the two components of the system impulse response in (2). The optical point spread function (PSF) is a convolving kernel with a sidelobe structure. For some imaging systems, this kernel is rotationally symmetric, although this is not the case for all systems. The Fourier transform of the PSF, the optical transfer function (OTF), is the self-correlation of the aperture weighting function for the system and as such has limited support and therefore limits the spatial frequencies in the image. The imaging sensor, on the other hand, has a spatial impulse response that is uniform over a certain rectangular region of the image formation plane. Its Fourier transform is therefore a two-dimensional sinc-cross-sinc function, which has a lobe structure, a network of line nulls and has its highest sidelobes on the principal coordinate axes. This frequency response function does not impose a hard limit on spatial frequencies, but it does fall off to very low values away from the frequency origin. Since the overall transfer function is the product of these two, when the optical PSF is smaller than the sensor impulse response, part of the sidelobe structure of the sensor frequency response function is within the frequency limit set by the OTF. Note that there is some sidelobe response all the way up to the spatial frequency limit imposed by the aperture, especially on the

principal axes. When multiple rotations exist in the component data images, the frequency domain below the OTF cutoff is “filled in” by the sidelobes of the sensor frequency response. It is well-known in the image restoration literature<sup>12</sup> that, when multiple images of the same source are blurred by different blurring kernels, a well-posed reconstruction problem can be formulated, even though individual images may constitute ill-posed reconstruction problems because of nulls in the frequency responses of their blurs. Naturally, the noise level of the component images will have the greatest effect on frequency components in the areas of smallest spatial frequency response, so rotations ensure that the highest possible SNR is available across all spatial frequencies, at least in a subset of the observed images.

Stark and Oskoui<sup>11</sup> proposed a reconstruction method for this problem based on the method of projections onto convex sets (POCS) in the spatial domain. In this paper, we will apply the POCS method in the frequency domain for reasons of computational efficiency. For the most part, frequency domain methods have been applied in the past to problems in which the component images were related by subpixel shifts, and not rotations. This is because rotations of a sampled image do not correspond to a simple operation in the frequency domain. We believe that the advantage of frequency domain projections is so great that even a relatively straight-forward approach involving a forward and reverse Fourier transform in every image projection yields an overall improvement in computational cost as compared with the spatial POCS approach.

As pointed out in two recent reviews of this area<sup>1,2</sup>, work on SR reconstruction began with the frequency domain method of Tsai and Huang<sup>13</sup>. This was followed by other frequency domain methods<sup>14,15</sup> which, like<sup>13</sup>, called for the solution of a large set of simultaneous equations representing the relationship between the component observed images and the final high-resolution image. Frequency domain methods were later superseded by spatial domain methods that provided more flexibility in the relationship between the component images and in the image blurring effects. Many approaches to spatial domain processing, such as regularized least-squares modeling and maximum a posteriori probability modeling have been reviewed in the literature<sup>1,2</sup>, but are beyond the present scope.

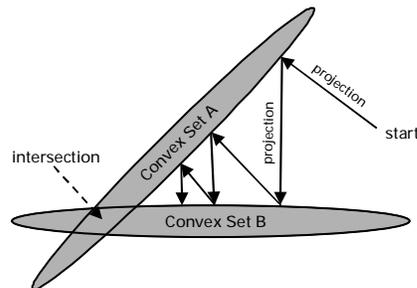


Figure 1: Projections onto convex sets.

POCS is a computational approach to finding an element of a feasible region defined by the intersection of a number of convex constraints, starting with an arbitrary infeasible point<sup>10,16</sup>. Figure 1 shows how convergence to the intersection is achieved by iterative projections onto the individual convex sets. This approach was first applied to the multi-image resolution enhancement problem by Stark and Oskoui<sup>11</sup> and their approach was later extended by Tekalp and others<sup>17,18</sup> to include motion blur (non-zero time aperture of the sensor) and non-rectangular sampling schemes. These methods reported the use of POCS in the spatial domain; in general, such algorithms suffer from slow convergence, but they are very flexible in the kinds of image degradations that can be addressed. The method of reference 19 is the only one known to us that applies POCS in the frequency domain. The conclusion of that work was that the frequency domain version of POCS seems to converge very quickly most of the time, but sometimes does not converge at all. The present authors interpreted this as a lack of any feasible region that satisfies all the constraints, and of course this is simply a matter of the size of the regions, since large enough regions will always intersect. The approach of reference 19 does not accommodate rotations.

## 2. FREQUENCY DOMAIN SR-POCS

The Frequency-Domain Super-Resolution by Projections Onto Convex Sets (FDSR-POCS) algorithm corrects the blurring caused by the image sensors and the associated aliasing in a single iterative computation. In FDSR-POCS, a high-resolution test image, referred to as the super-resolution image (SRI) is sequentially projected onto constraint sets constructed around observed component images called low-resolution images (LRI). Such constraint sets are composed of all high-resolution images that could have produced the low-resolution observation, after the effect of the PSF, imperfect sampling and aliasing. The projection operator identifies the feasible high-resolution image that is closest to the current SRI. This algorithm uses both the spatial and frequency domains, but does most of what would ordinarily be identified as projection in the frequency domain. For our purposes, an important benefit of POCS is that the running estimate is affected by only one observation at a time.

The projection operation is based on the relationship of a high-resolution image to a low-resolution one, and so it involves the smoothing effect of the image sensor, which is more efficiently computed in the frequency domain than in the spatial domain. Additionally, the aliasing relationship is very straightforward in the frequency domain. However, in order to address the rotation of the LRI's with respect to each other, which is not easily formulated in the frequency domain, FDSR-POCS uses resampling of the SRI in the spatial domain, which means that the SRI cannot be maintained solely in the frequency domain during the computation. Therefore we add to every image projection a reverse FFT, a spatial domain resampling operation and a forward FFT.

### 2.1 Structure of the FDSR-POCS algorithm

FDSR-POCS is designed to produce SR images from low-resolution images that are generated according to (3). Figure 2 shows an overview of the FDSR-POCS technique. In Figure 2, operations that are only carried out once are shown with dotted lines and arrows. These operations are depicted on the right-hand side, starting with the observed low-resolution images on the extreme right. For each of the  $M$  LRI's, a region of interest is defined, and the image is windowed so the region of interest is contained within in a reduced-size image at the sample rate of the LRI. The applied window is equal to unity over the region of interest and tapers off to zero outside that region. Such a window is meant to reduce ringing in the computed spectrum and to ensure that the region being enhanced is present in its entirety in each LRI. Each such windowed image depicts the same portion of the scene, but they are rotated with respect to each other and possibly shifted by a sub-pixel offset in an arbitrary direction. These images are given by the  $X_m(k_1, k_2)$  of (3). Each windowed image segment is then cropped to a smaller size and Fourier transformed using the FFT algorithm to a low-resolution spectrum (LRS). We denote this by

$$\tilde{X}_m(n_1, n_2) = \mathfrak{F}[X_m(k_1, k_2)] \quad \text{for } n_1, n_2 = 0, \dots, (K-1) \quad (5)$$

The super-resolution test image (SRI) is iteratively improved by projecting it onto constraint sets that are defined using the LRS's. Each projection onto a data-derived constraint set involves a resampling of the SRI to a super-resolution overlay image (SRO), an FFT, a projection of the Fourier-transformed SRO onto a convex set of SRO's that are consistent with an LRS, an inverse FFT and then a resampling back to the SRI coordinate system. If the projection operation resulted in no change in the SRO, then the inverse FFT and resampling operations can obviously be skipped. While this seldom happens with the point-by-point projection operator introduced first in this section, it does sometimes occur with the whole-image projection operator that is described later in the section. The resampling operation may be implemented by two-dimensional bandlimited resampling, since the frequency content of the image to be synthesized is known beforehand. If we are willing to pay the price of even greater oversampling (larger  $W$ ), then simpler interpolation methods, such as bilinear, bicubic or splines, can be used.

We apply rotations to the SRI, rather than trying to align the LRI as an initial step, because rotation of the highly sampled and unaliased SRI will result in less interpolation error than a similar operation applied to the LRI's.

While most of the constraint sets used by the algorithm may be defined relative to an LRS, some others may represent *a priori* knowledge of the scene or the imaging system. Such constraints are labeled "positivity" and "frequency limit" in Figure 2.

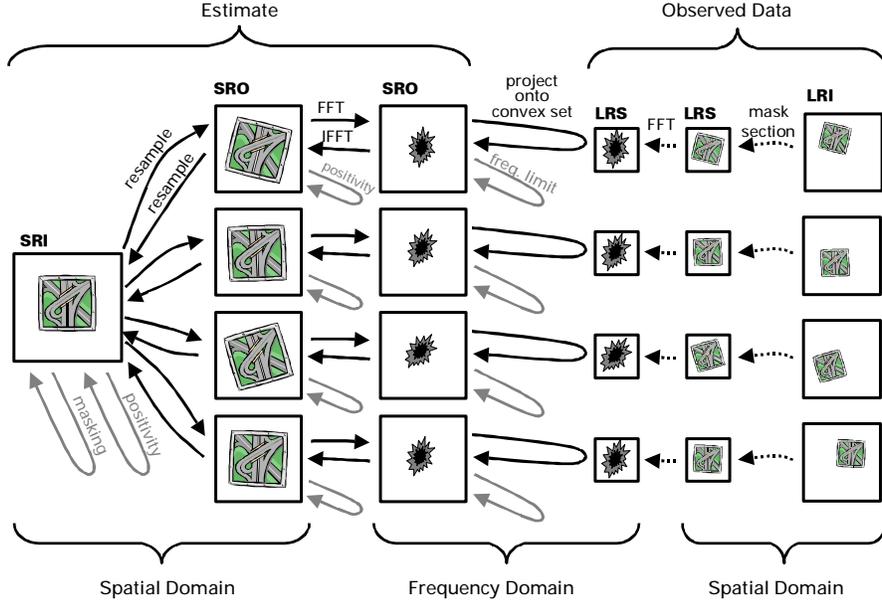


Figure 2: An overview of the FDSR-POCS algorithm, shown with four observed images.

The SRI is defined in its own coordinate system which is scaled so that the spatial sample rate of the SRI is  $W$  times that of the LRIs. The value of  $W$  is selected to be at least as large as the ratio of the width of the imaging PSF to that of the intensity sensor impulse response, rounded up to an integer up-sampling factor. The SRI has a high sample rate so that it can accurately represent the high spatial frequencies present in the synthesized image, given in (4), and so it can be resampled to implement rotation without significant distortion. We will denote the SRI as

$$I_{SRI}(m_1, m_2), \quad \text{for } m_1, m_2 = 0, \dots, (KW+D-1) \quad (6)$$

Note that the SRI is  $W$  times the size of the LRS in each dimension, plus some margin, given by  $D$ . This extra area is to allow for rotational re-sampling.

Each SRO coordinate system is defined relative to an LRS coordinate system in that they have the same rotation angle and subpixel shift relative to the SRI. An SRO coordinate system is simply that of the LRS scaled so that the sample rate in the SRO is  $W$  times the sample rate in the LRS. Thus the SRO is  $W$  times the size of the LRS in each dimension. This means that an image spectrum in the coordinate system of the SRO's can be easily related to an image spectrum in the associated LRS coordinate system in the Fourier domain. It is because we can form this relationship that we will repeatedly resample/transform the SRI into the SRO's and back. We denote the SRO associated with the  $m^{\text{th}}$  observed image as

$$\tilde{I}_{SRO,m}(r_1, r_2) = \mathfrak{F}[R_m(I_{SRI}(m_1, m_2))], \quad \text{for } r_1, r_2 = 0, \dots, (KW-1) \quad (7)$$

where  $R_m(\cdot)$  is an operator that resamples the SRI to rotate, shift and crop it so that it corresponds in coordinate axes to the  $m^{\text{th}}$  LRS.

The SRO and the LRS have the same origin. The SRO is scaled so that its sample rate is exactly  $W$  times greater than that of the LRS in each dimension. The SRO is the Fourier transform of an estimate of the scene portrayed by the  $m^{\text{th}}$  LRI, but with no image-plane sensor impulse response applied. The LRS is the DFT of the under-sampled observation of a portion of the scene, which includes the effect of the physical sensor. In the Fourier domain, we can transform the SRO into the LRS format by applying a weighting function to the SRO representing the frequency domain

action of the physical image sensor and then adding together sections of the SRO according to the aliasing principle. This relationship is given by

$$\tilde{I}_m(k_1, k_2) = \sum_{p=0}^{W-1} \sum_{q=0}^{W-1} F(k_1 + pK, k_2 + qK) \tilde{I}_{\text{SRO},m}(k_1 + pK, k_2 + qK) \quad (8)$$

where  $F(r_1, r_2)$  for  $r_1, r_2 = 0, \dots, (KW-1)$  is a frequency-domain weighting function which is the DFT of a spatial impulse response sampled at the SR rate, and the discrete frequency indices of the aliased result are in the range  $k_1, k_2 = 0, \dots, (K-1)$ .

The spatial impulse response that results in  $F(r_1, r_2)$  can be that of the physical intensity sensors, in which case the goal of the computation will be an image whose impulse response is the PSF of the imaging optics. It is also possible to synthesize the effect of another PSF, and to do so the weighting function must include a factor that converts the optical transfer function of the actual PSF into one corresponding to the desired PSF. For example, as mentioned above, the desired PSF could represent a uniform OTF, in which case the mainlobe resolution of the synthesized PSF would be higher than that of any PSF obtained through image formation alone, without postprocessing.

In the work that follows, it will be useful for us to write the aliasing rule of (8) in the form of an inner product. The rule says that the  $(k_1, k_2)^{\text{th}}$  component of the aliased and blurred version of the LRS is a linear combination of  $W^2$  terms of the SRO. It is only a matter of re-indexing the terms to obtain both the SRO components and the weights as vectors. We define the vector

$$\Phi(k_1, k_2) = \begin{bmatrix} F^*(k_1, k_2) \\ F^*(k_1, k_2 + K) \\ \vdots \\ F^*(k_1 + (W-1)K, k_2 + (W-1)K) \end{bmatrix} \quad (9a)$$

whose elements are complex conjugates of weights in (8), and also

$$\mathbf{X}_m(k_1, k_2) = \begin{bmatrix} \tilde{I}_{\text{SRO},m}(k_1, k_2) \\ \tilde{I}_{\text{SRO},m}(k_1, k_2 + K) \\ \vdots \\ \tilde{I}_{\text{SRO},m}(k_1 + (W-1)K, k_2 + (W-1)K) \end{bmatrix} \quad (9b)$$

Both definitions are for  $k_1, k_2 = 0, \dots, (K-1)$ . Note that the order of the components is not important as long as it is the same in both vectors. Now we can express (8) as an inner product

$$\tilde{I}_m(k_1, k_2) = \Phi^H(k_1, k_2) \mathbf{X}_m(k_1, k_2) \quad (10)$$

for  $k_1, k_2 = 0, \dots, (K-1)$ , and where the superscript H represents Hermitian transposition.

## 2.2 Sets of feasible SRO's

Feasible SRO's are those that, when transformed by way of (10), give a result that is close to the observed LRS. The notion of closeness is defined on a coefficient-by-coefficient basis in the LRS domain. Here is the definition of the vector of SRO DFT coefficients that are consistent with a single DFT coefficient in the observed LRS:

$$\left\{ \mathbf{X}_m(k_1, k_2) \mid \left| \Phi^H(k_1, k_2) \mathbf{X}_m(k_1, k_2) - \tilde{X}_m(k_1, k_2) \right|_2^2 \leq \epsilon^2 \right\} \quad (11)$$

This constraint set represents the constraints on a set of  $W^2$  DFT coefficients in that SRO that is rotated to match the  $m^{\text{th}}$  observed image. There are  $K^2$  such constraint sets, one for every DFT coefficient in the  $m^{\text{th}}$  LRS. Taken together, these constraints form a constraint set for the entire SRO.

In the  $W^2$ -dimensional space that represents the DFT coefficients that alias into a single DFT coefficient in the LRS, the inner product operation of (10) can be associated with projection onto a line. Thus the set described in (11) is the set of all points whose projections onto a line lie on a certain finite segment. The  $W^2$ -dimensional set described by (11) is finite in a single dimension, and is bounded by two  $(W^2-1)$ -dimensional hyperplanes. Thus the union of all  $K^2$  such constraints for the  $m^{\text{th}}$  image is finite in exactly  $K^2$  dimensions. Since there are  $(KW)^2$  pixels in the SRI, we expect to require at least  $W^2$  component images to define a bounded feasible region.

### 2.3 Computation of the projection of the SRO

The projection operation onto the set defined by (11) computes that element of the constraint set that is closest to the initial value of  $\mathbf{X}_m(k_1, k_2)$ . If the initial value is an element of the constraint set, it is left unchanged, otherwise the projector modifies the  $W^2$  components of  $\mathbf{X}_m(k_1, k_2)$  so that the qualification in (11) is satisfied with equality and the distance from the initial point to the computed point is minimized.

In this section we will derive a closed-form expression for the projector under the Euclidian norm. To simplify the notation, we will use  $\mathbf{b} = \tilde{\mathbf{X}}_m(k_1, k_2)$ ,  $\mathbf{X} = \mathbf{X}_m(k_1, k_2)$  and  $\Phi = \Phi(k_1, k_2)$ , with the understanding that the result applies to each DFT coefficient of each LRS, using the vector definitions of (9). We will assume that  $\mathbf{X}_0$  is the initial state of the SRO coefficients, and we will compute  $\mathbf{X}$ , the projection. It is quite easy to check to see if  $\mathbf{X}_0$  satisfies the constraint of (11) by computing  $|\Phi^H \mathbf{X}_0 - \mathbf{b}|^2$  and comparing it to  $\epsilon^2$ . If the initial value does not satisfy the constraint, the projection operation is performed.

This constrained minimization problem can be formulated using a Lagrange multiplier. We derive the projection operation by minimizing the following function for both  $\mathbf{X}$  and  $\lambda$ :

$$f(\mathbf{X}, \lambda) = |\mathbf{X} - \mathbf{X}_0|^2 + \lambda \left( |\Phi^H \mathbf{X} - \mathbf{b}|^2 - \epsilon^2 \right) \quad (12)$$

Setting the partial derivative with respect to  $\mathbf{X}$  equal to zero, we obtain

$$\mathbf{X} = (\mathbf{I} + \lambda \Phi \Phi^H)^{-1} (\mathbf{X}_0 + \lambda \mathbf{b} \Phi) \quad (13)$$

Now, using the small-rank adjustment lemma<sup>20</sup>, we find that

$$\mathbf{X} = \left[ \mathbf{I} - \frac{\lambda}{1 + \lambda |\Phi|^2} \Phi \Phi^H \right] (\mathbf{X}_0 + \lambda \mathbf{b} \Phi) \quad (14)$$

which we can rewrite as

$$\mathbf{X} = \mathbf{X}_0 - \frac{\lambda}{1 + \lambda |\Phi|^2} (\Phi^H \mathbf{X}_0 - \mathbf{b}) \Phi \quad (15)$$

This gives us the minimizing value of  $\mathbf{X}$  in terms of  $\lambda$ ; if we also find the minimizing  $\lambda$ , we can substitute it into (15) and so compute  $\mathbf{X}$ . We want a value of  $\mathbf{X}$  that is on the boundary of the constraint set, so we substitute (15) into the boundary condition  $|\Phi^H \mathbf{X} - \mathbf{b}|^2 = \epsilon^2$ . This gives us an equation that we can solve for lambda:

$$\left| \Phi^H \left[ \mathbf{X}_0 - \frac{\lambda}{1 + \lambda |\Phi|^2} (\Phi^H \mathbf{X}_0 - \mathbf{b}) \Phi \right] - \mathbf{b} \right|^2 = \epsilon^2 \quad (16)$$

After some algebra, this can be written as a quadratic equation in  $\lambda$ ,

$$|\Phi|^4 \lambda^2 + 2|\Phi|^2 \lambda + \left[ 1 - \frac{|\Phi^H \mathbf{X}_0 - \mathbf{b}|^2}{\varepsilon^2} \right] = 0 \quad (17)$$

and it can be solved using the quadratic formula. The roots of (17) are given by

$$\lambda_1, \lambda_2 = \frac{-1}{|\Phi|^2} \pm \frac{|\Phi^H \mathbf{X}_0 - \mathbf{b}|}{\varepsilon |\Phi|^2} \quad (18)$$

Both roots should be tried in (15). If we denote the root that minimizes  $|\mathbf{X} - \mathbf{X}_0|^2$  as  $\lambda_p$ , then the required projection is:

$$\mathbf{X} = \mathbf{X}_0 - \frac{\lambda_p}{1 + \lambda_p |\Phi|^2} (\Phi^H \mathbf{X}_0 - \mathbf{b}) \Phi \quad (19)$$

Note that we perform this computation only if  $|\Phi^H \mathbf{X}_0 - \mathbf{b}| > \varepsilon$ , so that the adjustment to  $\mathbf{X}_0$  in (19) is always computable. (If  $\Phi^H \mathbf{X}_0 = \mathbf{b}$ , then  $\lambda_p = \frac{-1}{|\Phi|^2}$ , which is a pole of the adjustment term in (19), so the condition ensures that the adjustment is finite.) Note also that  $|\Phi|^2$  can be pre-computed and applied to all M SRO's.

#### 2.4 Whole-image projection operator

Another approach to the projection operator can be taken. Rather than the point-by-point approach given above, we can use a whole-image approach, define a single constraint set for each observed low-resolution image, and use a projection operator that computes all of the DFT coefficients of the SRO in a single, iterative operation. Using the notation of Section 2.2, this single constraint set is defined by:

$$\left\{ \mathbf{X}_m \left| \sum_{k_1} \sum_{k_2} \left| \Phi^H(k_1, k_2) \mathbf{X}_m(k_1, k_2) - \tilde{X}_m(k_1, k_2) \right|_2^2 \leq \varepsilon^2 \right. \right\} \quad (20)$$

Note that, for i.i.d. additive Gaussian image noise this set is bounded by a level set of observation likelihood.

As in Section 2.1 we vectorize the 2D image operations so that we can use matrix and vector notation to derive the projection operation. In this case we will vectorize the operation that forms the entire LRS instead of a single coefficient of an LRS. We now define  $\Phi$  as the  $K^2$ -by- $(KW)^2$  matrix representing the operation given by (8),  $\mathbf{X}_0$  as the  $(KW)^2$ -by-1 vector composed of all the DFT coefficients in the initial, pre-projection value of the SRO,  $\mathbf{X}$  as the result of the projection, and  $\mathbf{b}$  as the  $K^2$ -by-1 vector of DFT coefficients of the LRS. The result derived will apply for each LRS. This vectorization is used in this derivation, but not in the execution of the resulting algorithm. The constraint set may be written using this notation as:

$$\left\{ \mathbf{X} \left| \left| \Phi^H \mathbf{X} - \mathbf{b} \right|_2^2 \leq \varepsilon^2 \right. \right\} \quad (21)$$

The constrained minimization problem that is the projection operation may be solved via a Lagrange multiplier. This function must be minimized over both  $\mathbf{X}$  and  $\lambda$ :

$$f(\mathbf{X}, \lambda) = |\mathbf{X} - \mathbf{X}_0|^2 + \lambda \left( \left| \Phi^H \mathbf{X} - \mathbf{b} \right|^2 - \varepsilon^2 \right) \quad (22)$$

Setting the partial derivative of  $f$ , with respect to  $\mathbf{X}$ , to 0 and solving yields:

$$\mathbf{X} = (\mathbf{I} + \lambda \Phi^H \Phi)^{-1} (\mathbf{X}_0 + \lambda \Phi^H \mathbf{b}) \quad (23)$$

Applying the small rank adjustment formula gives us:

$$\mathbf{X} = \left( \mathbf{I} - \lambda \Phi^H (\mathbf{I} + \lambda \Phi \Phi^H)^{-1} \Phi \right) (\mathbf{X}_0 + \lambda \Phi^H \mathbf{b}) \quad (24)$$

This solution involves a matrix inverse. Thanks to the frequency domain formulation for the problem at hand,  $\Phi$  has a specific structure that makes the  $K^2$ -by- $K^2$  matrix  $\Phi \Phi^H$  diagonal so this inverse is easy to compute. The final step of minimizing over  $\lambda$  requires a line search. The projection operator requires that the following equation be solved for  $\lambda$ :

$$\left\| \Phi^H \mathbf{X} - \mathbf{b} \right\|_2^2 = \varepsilon^2 \quad (25)$$

Equation (25) is solved using Brent's method<sup>21</sup>, although other approaches are possible. The value of  $\lambda$  found to satisfy (25) is used in (24) to compute the final result. The operations of (24) and (25) can be performed on data stored in memory-efficient image format. The data need not actually be reorganized into these large vectors and sparse matrices, which are used to derive the projection operator.

The whole-image projection operator given by (24) takes much longer to compute than does the point-by-point projection of the previous section, even accounting for the fact that the point-by-point projector must be performed  $K^2$  times to implement a single image projection. However, although the results of the algorithms using the two projectors seem similar to the eye in test cases run by the authors, they must be different solutions due to the differences in the constraint sets. We feel that this difference will probably manifest itself in the case of observed images with low SNR, since we feel that the shape of the constraints in the whole-image projection scheme will better match the shape of level sets of the noise distribution.

## 2.5 Structure of the whole-image observation convex constraint sets

An ellipsoidal convex set of acceptable SRO's for the whole-image projection exists in a high-dimensional space and is thus difficult to visualize. In this section we will focus in on two aspects of these sets: first eccentricity and then infinite extent, and discuss their implications on how the algorithm will progress.

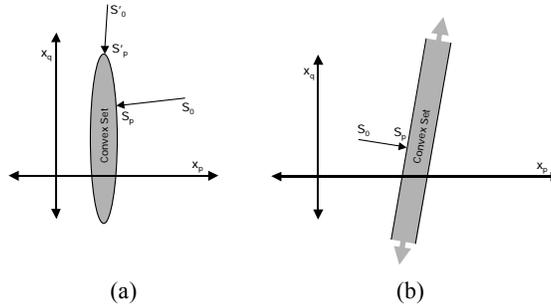


Figure 3: Convex constraint sets exhibit eccentricity and infinite extent.

Let us limit our view to two of the dimensions of a set of acceptable SRO's at a time and also artificially restrict our discussion to real numbers. Let  $x_p$  and  $x_q$  be 2 of the Fourier coefficients of the SRO that are not aliased with each other. Suppose  $x_p$  is a low frequency coefficient that is less affected by the PSF, and  $x_q$  is a high frequency coefficient that is more suppressed by the PSF. So, this particular observation provides more accurate information about  $x_p$  than about  $x_q$ .

So far as  $x_p$  and  $x_q$  are concerned, the defining equation for the set has the form:

$$\left\| \begin{bmatrix} \alpha_p x_p \\ \alpha_q x_q \end{bmatrix} - \begin{bmatrix} \beta_p \\ \beta_q \end{bmatrix} \right\|_2^2 \leq \varepsilon^2 \quad (26)$$

Here  $\alpha_p$  and  $\alpha_q$  are two particular coefficients of the PSF spectrum, with  $\alpha_p > \alpha_q$ , and  $\beta_p$  and  $\beta_q$  are two particular coefficients of the LRS spectrum. The resulting set is depicted in Figure 3a. Note how the set is narrow in the  $x_p$

dimension and wide in the  $x_q$  dimension. Because this observation provides more accurate information about  $x_p$ , the set tightly restricts the value of  $x_p$  and because this observation offers less accurate information about  $x_q$ , the set is more permissive in the values of  $x_q$  that are allowed.

In Figure 3a note the projection from  $S_0$  to  $S_p$ , which is the point within the convex set that is closest to  $S_0$ . In this projection,  $x_p$  changes a lot, and  $x_q$  changes a small amount. The change in  $x_p$ , where we have a more reliable observation, dominates, which is what we desire. Now consider the projection from  $S'_0$  to  $S'_p$ . At  $S'_0$ , the value for  $x_p$  is already consistent, so  $x_q$  will undergo more of a change. From most points in the vicinity of this set, the projection will alter  $x_p$  more than  $x_q$ . It is only when  $x_p$  is already consistent that  $x_q$  becomes more affected by the projection. Even though  $x_q$  cannot be determined accurately from this observation, projecting onto the consistency set for another observation may accurately determine  $x_q$ , depending on the various other PSF's. Through this mechanism applied simultaneously in all of the dimensions of the SRO, we expect the projections to first correct accurately observed, low spatial frequency, coefficients, and then to correct less accurately observed, higher spatial frequency, coefficients, to the extent possible.

The consistency sets are ellipsoids, but extend to infinity in many directions. As above, we again focus in on two of the dimensions,  $x_p$  and  $x_q$ . In this case, however, suppose that  $x_p$  and  $x_q$  are aliased with each other, and that  $x_p$  is a lower spatial frequency, better passed by the PSF. Now, so far as  $x_p$  and  $x_q$  are concerned, the defining equation for the set has the form,

$$\left| \alpha_p x_p + \alpha_q x_q - \beta \right|_2^2 \leq \epsilon^2 \quad (27)$$

Again,  $\alpha_p$  and  $\alpha_q$  are coefficients of the PSF spectrum, and  $\alpha_p > \alpha_q$ . The resulting set is depicted in Figure 3b. This infinite set is oriented nearly vertically in the figure because  $\alpha_p > \alpha_q$ . Consider what happens in the projection depicted in Figure 3b. To project from  $S_0$  to  $S_p$ , at the boundary of the consistency set,  $x_p$  is changed much more than  $x_q$  is changed.

When adjusting frequency coefficients that are confounded by aliasing, the inherent action is to favor adjusting the coefficient for the frequency that is better observed. This may be desirable because we can get from  $S_0$  to the consistency set by making a smaller, lower energy, change to  $x_p$ . This change may not be correct. It may be the case that to get close to the true scene we should be adjusting  $x_q$  more. Because of the aliasing there is no way to know from a single observation. However, the cumulative changes from projecting onto the different consistency sets from different observations should ultimately accurately determine  $x_p$  and  $x_q$ .

Each of these examples examines only 2 dimensions and presumes radically different PSF coefficients ( $\alpha_p$  and  $\alpha_q$ ) for demonstrative purposes. Consistency sets for real images are ellipsoids in very high dimensional spaces and the PSF coefficients vary over a wide range. To further complicate matters, all of the coefficients are actually complex. Needless to say, the true consistency sets are difficult to visualize. However, some of their characteristics are captured by these simple examples.

### 3. RESULTS

Figure 4 shows an example of the application of FDSR-POCS to a synthetic observed set of images. The images in 2a and 2b are 2 examples of the 16 total 128 by 128 pixel images that were generated synthetically from a high-resolution master with no image noise added. To make the reconstruction clearly visible in this example a radical oblong Gaussian PSF with  $\sigma=2.0$  pixels in the horizontal direction and  $\sigma=0.5$  pixels in the vertical direction. The 16 input images were generated with rotations distributed from  $0^\circ$  to  $180^\circ$ , but not at regular intervals, and small random sub-pixel shifts. Figure 4c shows the result of applying a Wiener filter to one of the input images. Because of the extreme blurring in the vertical direction the text on the plane wing is not legible in the enhanced image. Figure 4d shows the FDSR-POCS result where the text on the plane wing is now clearly legible. The frequency-domain weighting function,  $F(r_1, r_2)$ , used corresponds to a spatial domain PSF with  $\sigma=0.5$  input image pixels in each direction, and the expansion factor is  $W=2$ . Similar experiments with input images synthesized from a pseudo-noise master verify that reconstruction of spatial frequencies above the Nyquist rate of the input images can be achieved. In this figure, the input and Wiener filtered images are scaled by 2 so their size is the same as the FDSR-POCS image, which was created with scaling factor  $W=2$ .

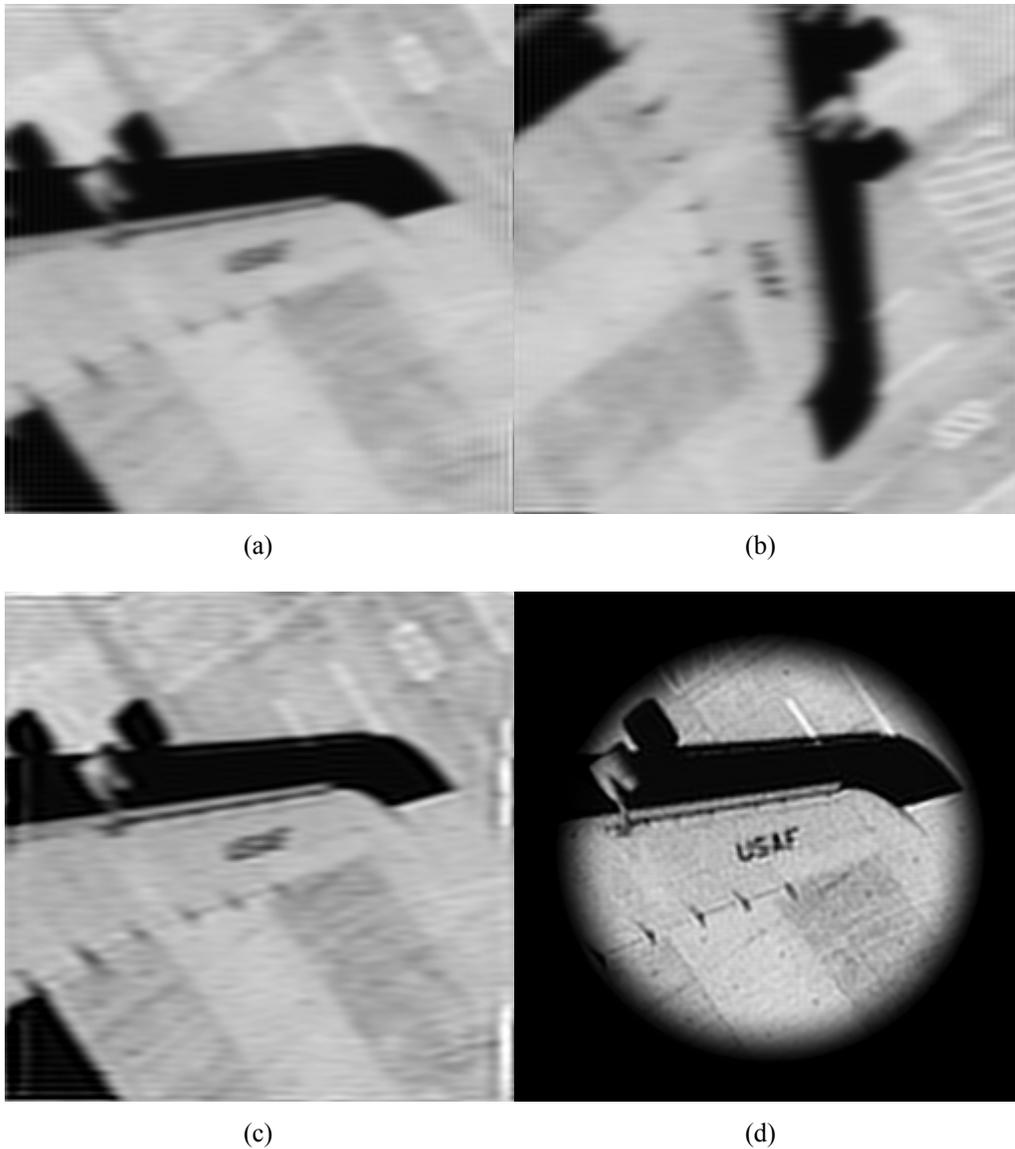


Figure 4: (a,b) 2 (of 16 total) synthetically generated input images; (c) Wiener filtered image; (d) FDSR-POCS result.

#### 4. DISCUSSION

In this paper we have proposed a frequency domain POCS algorithm for what we feel to be the canonical problem of super-resolution image synthesis. This algorithm is structured to accommodate rotations of the source relative to the imaging device, which we believe to help in producing a well-conditioned image synthesis problem.

Note that by defining the spatial frequency domain weighting function in the projection operator to correspond to the sensor impulse response, we would be asking the algorithm to synthesize an image in which all the frequencies passed by the system optics are weighted by the optical transfer function of the optics. This corresponds to an image with the natural PSF of the optics. Other point spread functions can be synthesized by using a weighting function in (8) that incorporates a weighting that re-shapes the OTF. The only constraint on this weighting is that it must have the same frequency domain support as the OTF.

## REFERENCES

1. S.C. Park, M.K. Park and M.G. Kang, "Super-resolution image reconstruction: a technical overview," *IEEE Signal Processing Magazine*, vol. 20, pp. 21-48, 2003.
2. S. Borman and R.L. Stevenson, "Super-resolution enhancement of low-resolution image sequences: a comprehensive review with directions for future research," Technical Report, Laboratory for Image and Signal Analysis, University of Notre Dame, July 1998.
3. Y.T. Lo, "On the theoretical limitation of a radio telescope in determining the sky temperature distribution," *J. Applied Physics*, vol. 32, pp. 2052-2054, Oct., 1961.
4. A.K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.
5. C.K. Rushforth and R.W. Harris, "Restoration, resolution and noise," *Journal of the Optical Society of America*, vol. 58, pp. 539-545, April, 1968.
6. A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Trans. Circuits and Systems*, vol. CAS-22, pp. 735-742, Sept., 1975.
7. G. Toraldo di Francia, "Degrees of freedom in an image," *Journal of the Optical Society of America*, vol. 59, pp. 799-804, July, 1969.
8. G.J. Buck and J.J. Gustincic, "Resolution limits of a finite aperture," *IEEE Trans. Antennas and Propagation*, vol. AP-15, pp. 376-381, May, 1967.
9. D.C. Youla and H. Webb, "Image restoration by the method of convex projections: part I – theory," *IEEE Trans. Medical Imaging*, vol. MI-1, pp. 81-94, Oct., 1982.
10. M.I. Sezan and H. Stark, "Image restoration by the method of convex projections: part II – applications and numerical results," *IEEE Trans. Medical Imaging*, vol. MI-1, pp. 95-101, Oct., 1982.
11. H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays using convex projections," *Journal of the Optical Society of America, Series A*, vol. 6, pp. 1715-1726, Nov., 1989.
12. D.C. Ghiglia, "Space-invariant deblurring given N independently blurred images of a common object," *Journal of the Optical Society of America, Series A*, vol. 1, pp. 1715-1726, April, 1984.
13. R.Y. Tsai and T.S. Huang, "Multiframe image registration and registration," *Advances in Computer Vision and Image Processing*, Ed. T.S. Huang, pp. 317-339, JAI Press, 1984.
14. S.P. Kim, N.K. Bose and H.M. Valenzuela, "Recursive reconstruction of high resolution image from noisy undersampled multiframes," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. 38, pp. 1013-1027, June, 1990.
15. S.P. Kim and W.-Y. Su, "Recursive high-resolution reconstruction of blurred multiframe images," *IEEE Trans. Image Processing*, vol. 2, pp. 534-539, Oct., 1993.
16. H.H. Bauschke and J.M. Borwein, "On projection algorithms for solving convex feasibility problems," *SIAM Review*, vol. 38, pp. 367-426, Sept., 1996.
17. A.M. Tekalp, M.K. Ozkan and M.I. Sezan, "High-resolution image reconstruction from lower-resolution image sequences and space-varying image restoration," *Proc. IEEE International Conf. on Acoustics, Speech and Signal Processing*, vol. 2, pp. 169-172, March, 1992.
18. A.J. Patti, M.I. Sezan and A.M. Tekalp, "Superresolution video reconstruction with arbitrary sampling lattices and nonzero aperture time," *IEEE Trans. Image Processing*, vol. 6, pp. 1064-1076, August, 1997.
19. H. Wei and T.D. Binnie, "High-resolution image reconstruction from multiple low-resolution images," *Proceedings of 7<sup>th</sup> International Conf. on Image Processing and its Applications*, vol. 2, pp. 596-600, July 1999.
20. R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
21. W.H. Press, S.A. Teukolsky, W.T. Vetterling and B.P. Flannery, *Numerical Recipes in C, 2d Ed.*, Cambridge University Press, 1992.