Motion-Free Superresolution

P.K. Vajapeyazula and A.N. Rajagopalan
Department of Electrical Engineering
Indian Institute of Technology – Madras
Chennai – 600 036, India
phanikv@lycos.com, raju@ee.iitm.ernet.in

Abstract

In this paper, we examine the theory of motion-free superresolution by formulating the problem in the Discrete Time Fourier Transform (DTFT) and the Discrete Fourier Transform (DFT) domain. Our approach provides some new insights into how aliasing and blurring can effect exact reconstruction of the superresolved image. For ease of understanding, the analysis, is initially carried out in the 1-D domain and then extended to the 2-D domain.

1. Introduction

Superresolution reconstruction methods attempt to recover a high resolution image from a set of low resolution observations. The idea of Superresolution from images with relative motion was first presented by Tsay and Huang [4]. Based on spatial aliasing effect, they used frequency domain approach to demonstrate the ability to reconstruct a high resolution image from several downsampled versions of it. A spatial domain alternative based on Papoulis [1] generalised sampling theorem was suggested by Ur and Gross [2]. In [3], Elad and Feuer have addressed the interesting problem of motion-free superresolution in which a high resolution image is derived from a set of blurred and downsampled versions of the original image.

A lower sampling rate implies a distortion of the image especially at edges. Blurring by different point spread function results in loss high frequency detail The aim of superresolution(SR), for the motion-free case, is to undo the effects of blurring and aliasing, by making use of the information in the given set of observations. Although different algorithms have been proposed in the literature for motion-free SR, a signal processing perspective of the problem is lacking. Our effort in this paper is directed towards this end. We assume that there is no relative motion among the different observations available. Specifically, we examine the case of *exact* reconstruction. Our analysis leads to a set of (natural) conditions whence motion-free SR is **not** possible. It explains the appropriateness of choice of blur and importance of number of observations. For exact reconstruction, for uniform downsampling by factor of *M*, we need atleast *M different* observations. The blur kernels, which are low-pass filters must span the *entire* frequency range of the original image. Importantly, the effects of downsampling can be reversed, only if the information lost is in the form of aliasing , and not while blurring.

2. Formulation in the 1-D domain

Let the original sequence $x[n]$ of length N have a DTFT $X(e^{j\omega})$ and a frequency spread $(-\omega_N, \omega_N)$. Let there be a set of *M* sequences, $h_k[n]$, also of length N, with DTFT $H_k(e^{j\omega})$ which vanishes outside $(-\omega_k, \omega_k)$. The k^{th} observation $y_k[n]$ is obtained by blurring $x[n]$ with sequence $h_k[n]$ and then downsampling by a factor of M.

2.1. DTFT Analysis

The, DTFT of $y_k[n]$ can then be expressed as

$$Y_k(e^{j\omega}) = \frac{1}{M}\sum_{i=0}^{M-1} X(e^{j(\frac{\omega-2\pi i}{M})})H_k(e^{j(\frac{\omega-2\pi i}{M})}) \qquad k = 0, 1 \ldots M-1 \qquad (1)$$

The above equation can be interpreted as follows. The spectrums of $x[n]$ and $h_k[n]$ are each scaled up by M. Then, M versions (each shifted by 2π) of their product, are added up to generate the output $Y_k(e^{j\omega})$. Clearly, when there is no aliasing of the shifted spectrums, the use of an appropriate low pass filter would give us all the information needed to reconstruct $x[n]$ exactly, and just one observation would suffice. This is, ofcourse, assuming that $H(e^{j\omega})$ doesn't vanish in the range $(-\omega_N, \omega_N)$. However, when there is aliasing in all the M observations (leading to loss of frequency content) there is a need for multiple observations. Using the matrix-vector notation, equation (1) can be rearranged as

$$\underline{Y} = \underline{H}\,\underline{X} \qquad (2)$$

where

$$\underline{Y} = \begin{bmatrix} Y_1(e^{j\omega}) \\ Y_2(e^{j\omega}) \\ \vdots \\ Y_M(e^{j\omega}) \end{bmatrix}, \underline{X} = \begin{bmatrix} X(e^{j\frac{\omega}{M}}) \\ X(e^{j\frac{\omega-2\pi}{M}}) \\ \vdots \\ X(e^{j\frac{\omega-2\pi(M-1)}{M}}) \end{bmatrix}$$

and

$$\underline{H} = \begin{bmatrix} H_1(e^{j\frac{\omega}{M}}) & H_1(e^{j\frac{\omega-2\pi}{M}}) & \dots & H_1(e^{j\frac{\omega-2\pi(M-1)}{M}}) \\ H_2(e^{j\frac{\omega}{M}}) & H_2(e^{j\frac{\omega-2\pi}{M}}) & \dots & H_2(e^{j\frac{\omega-2\pi(M-1)}{M}}) \\ \vdots & \vdots & & \vdots \\ H_M(e^{j\frac{\omega}{M}}) & H_M(e^{j\frac{\omega-2\pi}{M}}) & \dots & H_M(e^{j\frac{\omega-2\pi(M-1)}{M}}) \end{bmatrix}$$

This $\underline{X}$ can be reconstructed only if the matrix $\underline{H}$ is non-singular. It may be noted that each of the terms in the vector $\underline{X}$ only gives a part of the entire frequency spectrum. For example, for $\omega \in (-\pi, \pi)$, the term $X(e^{j\frac{\omega-2\pi i}{M}})$ generates the frequency spectrum, in the range $(\frac{-\pi-2\pi i}{M}, \frac{\pi-2\pi i}{M})$. Thus, each term only generates $\frac{1}{M}^{th}$ of the entire frequency spectrum of $x[n]$, and hence, there is the need of *M* observations to obtain $x[n]$.

A completely trivial case, where the matrix $\underline{H}$ turns out to be singular is when 2 or more blur sequences are exactly identical which renders motion-free SR impossible. Hence, we need *M* independent observations. What may not be entirely obvious is that SR reconstruction is not possible if none of the blur sequences span any particular frequency range of $X(e^{j\omega})$. For example, take the case where

$$\omega_k < \omega_N, \quad \forall\, k \in (1, M)$$

Surely, $\underline{H}$ turns out to be singular. The above condition translates to loss of frequency content during a low-pass blurring operation. Exact reconstruction methods can only reverse the effects of aliasing. If information is lost over any interval within $(-\omega_N, \omega_N)$, during blurring, the original sequence $x[n]$ cannot be regenerated. Ofcourse, the blur kernels can take any values outside the interval $(-\omega_N, \omega_N)$.

The problem is further analysed, using the Discrete Fourier Transform (DFT) which would give further insight into the problem and also, a practical method to check if exact reconstruction is possible.

2.2. DFT Analysis

The DFT of a sequence of length N can be obtained from its DTFT by sampling it at N points. Now, $x[n]$ is an N-length sequence which has been blurred and downsampled by a factor of M, to get the observations. We sample the DTFT of the l^{th} observation $y_k[n]$ at $\frac{N}{M}$ points to obtain from equation (1)

$$Y_l[k] = \frac{1}{M} \sum_{i=0}^{M-1} X(e^{j(\frac{2\pi k}{N} - \frac{2\pi i}{M})}) H_l(e^{j(\frac{2\pi k}{N} - \frac{2\pi i}{M})}), \quad (3)$$

$$k \;=\; 0, 1, \dots (\tfrac{N}{M}) - 1$$

The above equation relates the DFT of the l^{th} observation to the DFT of the original sequence and the DFT of the l^{th} blur kernel. Re-arranging the equation (3) in the matrix-vector form, we get

$$\underline{Y_l} = \underline{H_l}\ \underline{X} \quad (4)$$

where

$$\underline{Y_l} = \begin{bmatrix} Y_l[0] \\ Y_l[1] \\ \vdots \\ Y_l[\frac{N}{M} - 1] \end{bmatrix} \qquad \underline{X} = \begin{bmatrix} X[0] \\ X[1] \\ \vdots \\ X[N-1] \end{bmatrix}$$

and $\underline{H_l} =$

$$\begin{bmatrix} H_l[0] & 0 & \dots & 0 & H_l[\frac{N}{M}] & 0 & \dots & \dots & 0 & H_l[\frac{\overline{M-1}N}{M}] & 0 & \dots \\ 0 & H_l[1] & 0 & \dots & 0 & H_l[\frac{N}{M}+1] & 0 & \dots & \dots & 0 & H_l[\frac{\overline{M-1}N}{M}+1] & 0 \\ 0 & 0 & \ddots & 0 & \dots & 0 & \ddots & 0 & \dots & 0 & \ddots & 0 \\ 0 & \dots & 0 & H_l[\frac{N}{M}-1] & 0 & \dots & 0 & H_l[\frac{2N}{M}-1] & 0 & \dots & 0 & H_l[N-1] \end{bmatrix}$$

From $\underline{Y_l}$ (l = 1,2 ... M) the *M* equations obtained can be stacked up as

$$\underline{Y} = \underline{H}\ \underline{X} \quad (5)$$

where

$$\underline{Y} = \begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_M \end{bmatrix} \quad \text{and} \quad \underline{\mathbf{H}} = \begin{bmatrix} \underline{H}_1 \\ \underline{H}_2 \\ \vdots \\ \underline{H}_M \end{bmatrix}$$

Clearly, exact reconstruction is possible only if the matrix $\underline{H}$ in equation (5) is invertible. Again, when any two blur sequences are exactly identical, the matrix $\underline{H}$ becomes singular which asserts the need for *M* independent observations.

Yet, another case where the matrix $\underline{H}$ turns out to be singular is easily identified as shown below.

$$H_l[0], \quad H_l[\frac{N}{M}], \quad \dots \quad H_l[\frac{\overline{M-1}N}{M}] \;=\; 0$$

for any one of the M observations.

In earlier papers on motion-free SR, the relation between $x[n]$ and the blurred and downsampled k^{th} observation $y_k[n]$ has been expressed in the spatial domain as

$$\underline{y_k} \;=\; \underline{D_k}\ \underline{h_k}\ \underline{x}$$

where $\underline{x}$ and $\underline{y_k}$ are lexicographically obtained from $x[n]$ and $y_k[n]$, $\underline{D_k}$ is the downsampling matrix, while $\underline{h_k}$ is the blur matrix obtained from the blur sequence $h_k[n]$.

An analysis in the frequency domain would follow from this equation by multiplying both sides by DFT matrices of appropriate dimensions (note that the matrix $\underline{D_k}$ will not be diagonalised by the DFT matrices). The resultant system of equations will be identical to what we obtained in equations (4) and (5). However, the process by which we arrive at equation (5) throws much more light and gives new insights into the motion-free SR problem.

3. Formulation in 2-D domain

The formulation in the 2-D domain follows in a rather straight forward manner. Let $x[m,n]$ be an image of size $N\,X\,N$ with the Fourier Transform given by $X(\omega_1,\omega_2)$. Further, let there be M^2 blurring kernels, with point spread function of the l^{th} kernel being $h_l[m,n]$ with the Fourier Transform $H_l(\omega_1,\omega_2)$. Then the l^{th} observation $y_l[m,n]$ is obtained by blurring $x[m,n]$ with $h_l[m,n]$ and then downsampling the resultant image by a factor of M.

3.1. DTFT Analysis

The Fourier Transform $Y_l(\omega_1,\omega_2)$ of the l^{th} observation is given by equation (6) Similar to the 1-D case,we observe that these set of M^2 equations can be solved for obtaining the entire frequency spectrum of $X(\omega_1,\omega_2)$ if we have M^2 independent observations. These equation can be arranged in the matrix-vector notation to obtain

$$\underline{Y} = \underline{H}\,\underline{X}$$

where the $(i,j)^{th}$ element of $\underline{H}$ is given by

$$H(i,j) = H_i(\frac{\omega_1 - m\pi}{M}, \frac{\omega_2 - n\pi}{M})$$

where, $m = (j \mod M)$, and $n = int\ (j/M)$

$$\underline{X} = \begin{bmatrix} X(\frac{\omega_1}{M}, \frac{\omega_2}{M}) \\ X(\frac{\omega_1-2\pi}{M}, \frac{\omega_2}{M}) \\ \vdots \\ X(\frac{\omega_1-(M-1)\pi}{M}, \frac{\omega_2}{M}) \\ X(\frac{\omega_1}{M}, \frac{\omega_2-2\pi}{M}) \\ X(\frac{\omega_1-2\pi}{M}, \frac{\omega_2-2\pi}{M}) \\ \vdots \\ X(\frac{\omega_1-(M-1)\pi}{M}, \frac{\omega_2-2\pi}{M}) \\ \vdots \\ X(\frac{\omega_1}{M}, \frac{\omega_2-(M-1)\pi}{M}) \\ X(\frac{\omega_1-2\pi}{M}, \frac{\omega_2-(M-1)\pi}{M}) \\ \vdots \\ X(\frac{\omega_1-(M-1)\pi}{M}, \frac{\omega_2-(M-1)\pi}{M}) \end{bmatrix} \quad \underline{Y} = \begin{bmatrix} Y_1(\omega_1,\omega_2) \\ Y_2(\omega_1,\omega_2) \\ \vdots \\ Y_l(\omega_1,\omega_2) \\ \vdots \\ Y_{M^2}(\omega_1,\omega_2) \end{bmatrix}$$

So, the **H** matrix thus obtained, will have to be invertible for Motion-Free SR to be possible. And, in the case where any 2 observations become exactly identical, this is clearly not possible. The need for M^2 observations is hence justified.

3.2. DFT Analysis

Once again, using a similar analysis as in the 1-D case, we sample the Fourier Transform of the l^{th} observation (while noting that it is an image of size $N/M\,X\,N/M$) appropriately to obtain equation (7). Equation (7) can be rewritten to obtain the relation between the DFT of the $y_l[m,n]$ and the DFTs of $x[m,n]$ and $h_l[m,n]$ as shown in equation (8)

It is observable from equation (8) that each term in the expression for $Y_l(k_1,k_2)$ comprises of M^2 terms from $X(k_1,k_2)$. Each observation $Y_k(k_1,k_2)$ leads to $\frac{N^2}{M^2}$ equations and hence, in case we have M^2 independent observations, we can solve for the N^2 DFT co-efficients of $X(k_1,k_2)$ and hence reconstruct the original image $x[m,n]$.

$$Y_l(\omega_1,\omega_2) = \frac{1}{M^2}\sum_{p=0}^{M-1}\sum_{q=0}^{M-1} X(\frac{\omega_1-2\pi p}{M}, \frac{\omega_2-2\pi q}{M})\, H_l(\frac{\omega_1-2\pi p}{M}, \frac{\omega_2-2\pi q}{M}) \tag{6}$$

$$Y_l(k_1,k_2) = \frac{1}{M^2}\sum_{p=0}^{M-1}\sum_{q=0}^{M-1} X(\frac{2\pi k_1}{N}-\frac{2\pi p}{M}, \frac{2\pi k_2}{M}-\frac{2\pi q}{M})\, H_l(\frac{2\pi k_1}{N}-\frac{2\pi p}{M}, \frac{2\pi k_2}{N}-\frac{2\pi q}{M}) \tag{7}$$

$$Y_l(k_1,k_2) = \frac{1}{M^2}\sum_{p=0}^{M-1}\sum_{q=0}^{M-1} X((k_1-\frac{pN}{M})_N, (k_2-\frac{qN}{M})_N)\, H_l((k_1-\frac{pN}{M})_N, (k_2-\frac{qN}{M})_N) \tag{8}$$

These equations can be re-arranged, using the matrix-vector notation to obtain

$$\underline{Y_l} = \underline{H_l}\,\underline{X}$$

where

$$\underline{Y_l} = \begin{bmatrix} Y_l[0,0] \\ Y_l[0,1] \\ \vdots \\ Y_l[0,\frac{N}{M}-1] \\ Y_l[1,0] \\ Y_l[1,1] \\ \vdots \\ Y_l[1,\frac{N}{M}-1] \\ \vdots \\ Y_l[\frac{N}{M}-1,0] \\ Y_l[\frac{N}{M}-1,1] \\ \vdots \\ Y_l[\frac{N}{M}-1,\frac{N}{M}-1] \end{bmatrix} \quad \underline{X} = \begin{bmatrix} X[0,0] \\ X[0,1] \\ \vdots \\ X[0,\frac{N}{M}-1] \\ X[1,0] \\ X[1,1] \\ \vdots \\ X[1,\frac{N}{M}-1] \\ \vdots \\ X[\frac{N}{M}-1,0] \\ X[\frac{N}{M}-1,1] \\ \vdots \\ X[\frac{N}{M}-1,\frac{N}{M}-1] \end{bmatrix}$$

And the $\underline{H_l}$ matrix is defined using its $(i,j)^{th}$ element as shown below. The elements of the i^{th} row contribute to the formation of $Y_l[m,n]$ where,

$m = int(\frac{i}{(N/M)})$ and $n = i\,mod(\frac{N}{M})$

$H_l(i,j) = 0$, except for M^2 locations in each row, where the locations are given by

$$H_l(i,j) = hl_{dft}[x,y], \quad x = (m - p\frac{N}{M}),\ y = (n - q\frac{N}{M})$$

$p, q = 0, 1 \ldots (M-1)$ and $j = x\frac{N}{M} + y$

where $hl_{dft}[x,y]$ gives the dft of the l^{th} blur kernel. The M^2 equations thus obtained are stacked up to obtain the familiar equation

$$\underline{Y} = \underline{H}\,\underline{X}$$

As in the earlier case, when two observations are exactly identical, the SR reconstruction is not possible. It is also expected that, as the blur kernels become sufficiently identical to each other, the $\underline{H}$ matrix becomes more and more ill-conditioned and hence, the quality of the reconstructed image goes down.

4. Increase/Decrease in number of observations

It is interesting to analyse the case where there are more than adequate number of observations available (more than M^2 for the case of exact reconstruction). It is intutively expected that more number of observations would provide additional information and hence there would be an improvement in the quality of the image obtained. But this is not the case always as we shall see shortly. A Least squares estimate of the image is obtained in case we have more than adequate number of images. The equation giving the least square estimate is

$$\mathbf{R}\underline{X}_{LS} = \underline{\mathbf{P}}$$

where

$$\mathbf{R} = \mathbf{H}^T\mathbf{H} \quad \text{and} \quad \underline{\mathbf{P}} = \mathbf{H}^T\underline{\mathbf{Y}}$$

So in this case the ill-condition-ness of $\mathbf{R}$ is to be discussed. It can be seen that, in the worst case where two rows are exactly identical in $\mathbf{H}$, $\mathbf{R}$ would turn out to be singular. So in the case where any two observations are sufficiently close to each other, $\mathbf{R}$ turns out to be ill-conditioned and the quality of image obtained goes down. So, the mere presence of an extra observation does not help but it is rather how *different* it is, is what matters.

Similarly we now analyse the case where we have less than M^2 observations. Using, a min-norm solution for this method, we recover the original image using the following equations.

$$\underline{X}_{MN} = \mathbf{H}^T(\mathbf{H}\mathbf{H}^T)^{-1}\underline{\mathbf{Y}}$$

Once again, as the observations come closer, the quality of the reconstructed image becomes worse.

The common observation from all these cases, is that the more than just the number of observations, the quality of the various observations also matters and a judicious choice of blur kernels is necessary for the extraction of a high resolution image.

5. Simulation Results

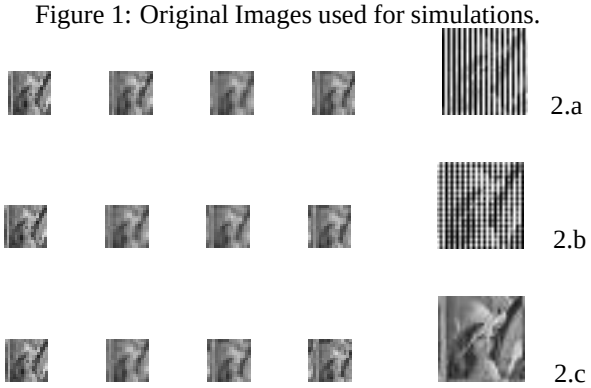
The blurring kernels used for this purpose are Gaussian filters with blur parameter σ_i for the i^{th} kernel. The original pictures are of size 32X32 and the observations have been uniformly downsampled by a factor of 2. The simulations are mainly directed at showing how the quality of the reconstructed image goes down, as the observations become more and more simialr to one another. The simulations also show how the variation in the choice of blur kernels reflects in the ill-condition-ness of the matrix to be inverted ($\mathbf{H}$ in the case of exact reconstruction , $\mathbf{H}^T\mathbf{H}$ in the case of Least-squares estimate and $\mathbf{H}\mathbf{H}^T$ in the case of Min-norm solutions.)

5.1. Exact reconstruction

Since the downsampling factor is 2, a total of 4 images are needed for exact reconstruction.

lena sky scene eye

Figure 1: Original Images used for simulations.



2.a
2.b
2.c

Figure 2: Observations and the reconstructed image
2.a $\sigma_1 = 0.2, \sigma_2 = 0.45, \sigma_3 = 0.7, \sigma_4 = 0.95$
2.b $\sigma_1 = 0.2, \sigma_2 = 0.5, \sigma_3 = 0.8, \sigma_4 = 1.1$
2.c $\sigma_1 = 0.2, \sigma_2 = 0.6, \sigma_3 = 1.0, \sigma_4 = 1.4$

a b c

Figure 3: Reconstruction with 3 sets of σs described in 2a, 2b and 2c.

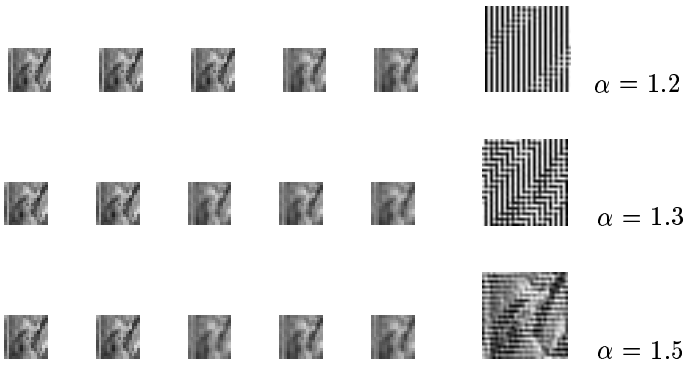
a b c

Figure 4: Reconstruction of 'scene' with 3 sets of σs described in 2a, 2b and 2c.

5.2. More Number of observations - Least Squares Estimate

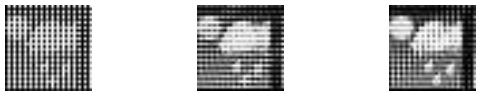
A test run with 5 observations is given below.The variances for the 5 observations are controlled with a parameter α and

$$\sigma_1 = 0.2, \sigma_2 = \alpha\sigma_1, \sigma_3 = \alpha\sigma_2, \sigma_4 = \alpha\sigma_3, \sigma_5 = \alpha\sigma_4$$



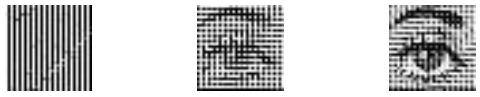
$\alpha = 1.2$
$\alpha = 1.3$
$\alpha = 1.5$

Figure 5: Observations and the reconstructed image



$\alpha = 1.2$ $\alpha = 1.3$ $\alpha = 1.5$

Figure 6: Reconstruction of 'sky' with 5 observations.

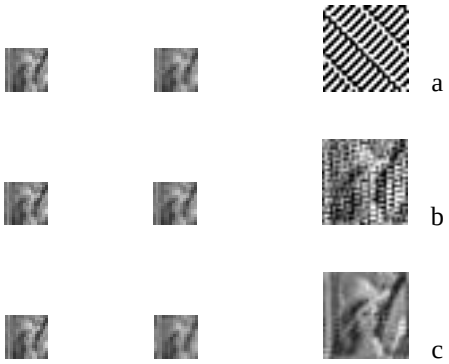


$\alpha = 1.2$ $\alpha = 1.3$ $\alpha = 1.5$

Figure 7: Reconstruction of 'eye' with 5 observations.

5.3. Less Number of observations - Min-norm Solution

A test run with 2 observations is given below.
$\sigma_1 = 0.4, \sigma_2 = 0.5$ for observation a
$\sigma_1 = 0.4, \sigma_2 = 0.6$ for observation b
$\sigma_1 = 0.4, \sigma_2 = 0.7$ for observation c



a
b
c

Figure 8: Observations and the reconstructed image

a b c

Figure 9: Reconstruction of 'sky' with 2 observations.

a b c

Figure 10: Reconstruction of 'eye' with 2 observations.

5.4. Reciprocal Condition Estimator

These graphs show how the ill-conditionness increases as observations come closer. The matrices considered are $\mathbf{H}$ for exact reconstruction, $\mathbf{H^T H}$ for Least-squares estimate and $\mathbf{H H^T}$ for Min-norm solution.

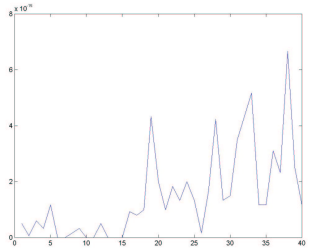


Figure 11: Plot of Reciprocal Condition Estimator with varying $\alpha = x/10$ (Exact Reconstruction) $(\sigma_1 = \alpha, \sigma_2 = 2\alpha, \sigma_3 = 3\alpha, \sigma_4 = 4\alpha)$

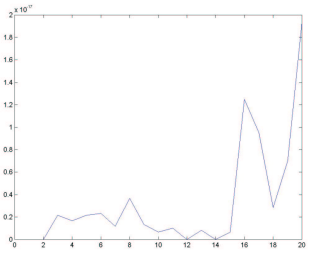


Figure 12: Plot of Reciprocal Condition Estimator with varying $\alpha = 1 + x/20$(Least Squares Estimate)

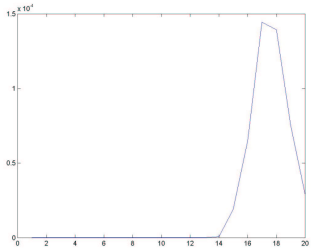


Figure 13: Plot of Reciprocal Condition Estimator with varying $\alpha = 1 + x/20$ (Min-norm solution

6. Conclusions

The case of Motion Free Superresolution has been explored in this paper. An analysis was carried out in the DTFT and the DFT domains and it was observed that exact reconstruction would be possible only when there are atleast M^2 observations. It was also analysed that as the observations come "closer" to each other, the quality of the reconstructed image deteriorates. A similar analysis was carried out in the case where the number of observations was not equal to M^2. A Least Squares estimate was used for the case of greater number of observations and a Min-norm solution was used for the case with lesser number of observations. These too, when analysed, show that the quality of the reconstructed image becomes worse as the observations become sufficiently identical.

References

[1] A.Papoulis. Generalized Sampling Theorem. *IEEE Transactions Circuits Syst*, CAS-24:652–654, November 1977.

[2] H.Ur and D.Gross. Improved Resolution from Sub-pixel Shifted Pictures. *CVGIP : Graph,Models, Image Processing*, 54:181–186, March 1992.

[3] M.Elad and A.Feuer. Restoration of a Single Superresolution Image from Several Blurred,Noisy and Undersampled Measured Images. *IEEE Transactions on Image Processing*, 6(12):1646–1658, December 1997.

[4] T.S.Huang and R.Y.Tsay. Multiple Frame Image Resoration and Registration. *Advances in Computer Vision and Image Procesing, JAI Press Inc.*, vol. 1:317–339, 1984.