EXTENDING THE DEPTH OF FOCUS USING DIGITAL IMAGE FILTERING

by

GUANG-HUA HU

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Electrical Engineering

APPROVED:

Dr. R.J. Vieper, Chairman

Dr. T.-C. Poon

Dr. R.O. Claus

April, 1987

Blacksburg, Virginia

EXTENDING THE DEPTH OF FOCUS USING DIGITAL IMAGE FILTERING

by

GUANG-HUA HU

Dr. R.J. Pieper, Chairman

Electrical Engineering

(ABSTRACT)

Two types of image processing methods capable of forming a composite image from a set of image slices which have in-focus as well as out-of-focus segments are discussed.

The first type is based on space domain operations and has been discussed in the literature. The second type, to be introduced, is based on the intuitive concept that the spectral energy distribution of a focused object is biased towards lower frequencies after blurring. This approach requires digital image filtering in the spatial frequency domain.

A comparision among methods of both types is made using a quantitative fidelity criterion.

Acknowledgements

The author wishes to sincerely thank Dr. R.J. Pieper for serving as his graduate advisor and chairman of his graduate committee.

A thank you also goes to Dr.T.-C. Poon and Dr.R.O. Claus for serving as his committee members.

The author has deep-hearted appreciation for his dear aunt who gave him the chance to pursue a higher degree.

A special thanks goes to his wife, , for her support and understanding.

Table of Contents

List of Illustrations

List of Tables

Chapter One Introduction

Digital image processing has been developed in response to three major problems. They are picture digitization and coding; picture enhancement and restoration; picture segmentation and description. The purpose of picture enhancement and restoration is to improve the degraded pictures. One type of degradation involves blur which is a form of spatial degradation.

One area of image restoration is extending depth of focus. In order to get a good image of any object, we have to put the object in focus. Of course, as long as we put the object within a certain range of the in focus plane, the image of object is, by definition of the depth of focus, acceptable. This range depends on physical considerations discussed in Appendix A. It is shown in the appendix the wider the aperture of an imaging system, the lower is the depth of field and higher the resolution. Increasing the resolution through an increase in the aperture size will lend to a corresponding decrease in the depth of field. That is why the increase of the depth of focus is of interest, especially in high resolution or high aperture image systems.

Various techniques have been proposed for extending the depth of field. All techniques which we discuss entail digitally analyzing a sequence of images each of which is

composed of in-focus as well as out-of-focus regions. Previous methods are implemented without translating the spatial domain information to spatial frequency domain. In this thesis, the author provides several new methods which operate on the spatial frequency domain of the image. Then the comparison is made among these spatial frequency domain methods and previous spatial domain methods based on RMS values.

In chapter two, there is a brief review of literature. Chapter three and four represents the main part of this thesis. In chapter three, four new methods are introduced for extending the depth of field. We consider all methods discussed in this chapter of the same type, since they are performed in the spatial frequency domain. In chapter four, a total of four methods implemented in the spatial domain are performed. The algorithms in chapter four are based on previous work. The main difference between the methods in this chapter and in the previous chapter is that all operations are done in the spatial domain. Chapter five is devoted to a comparison of all the methods. Finally, chapter six presents the conclusion. There are four appendices in this thesis. Appendix A, B, C describe the necessary background, tools and system. In detail, appendix A explains the relationship among the aperture of lens, depth of focus and resolution in an optical system. Appendix B is devoted to the digital image processing system (GIPSY) and appendix C to the fast Fourier transform. The main content of chapter three has already been concentrated in preprint form, which has been presented at IEEE, SSST, March 1987. This paper is in appendix D for reference.

Chapter Two Review of the Literature

There are several techniques which have shown improvement in the optical depth of focus. In this chapter, we briefly survey the literature.

M. Mino and Y. Okono proposed a method to improve the depth by using shaded aperture.[1] Near focus, the optical transfer function (OTF) for the shaded aperture has higher values in the low frequency region than the other aperture. Since the low frequency components of an object play the main role in image formation, shaded aperture improves the defocused image. They found that any shaded aperture, as long as it's OTF, is a monotonically decreasing nonnegative function, would improve the defocused image.

J. T. McCrickerd developed another technique to increase the depth of focus by using a high aspect ratio annular aperture combined with an optical equalization of the modulation transfer function (MTF).[2] The MTF of annular aperture is only large at the low spatial frequency, because of the diffraction limitations. The author used a spatial equalization filter in coherent system to overcome these limitations. The spatial equlization filter was located in the Fourier transform plane. The result demonstrated an extension of depth of focus.

G. Hausler on his own and with E. Korner provided basically two types of techniques to improve the depth of field.[3][4] The first technique involved two steps. The pure superposition of successively focussed images improve the quality of the defocussed parts, but at the same time, it caused the sharp part to become degraded. To overcome this difficulty, the author first did the incoherent superposition by moving a microscopic object along the optical axis; then they did coherent filtering by deconvolution. The second method is based on following idea. Since the high aperture imaging of three-dimensional objects which only one plane can be sharply imaged simultaneously, The author only selected the sharp areas of each image, then put these sharp image pieces together to form the result image. The key was how to select the sharp part of each image slice. The author provided two criterions for sharpness in spatial domain. One was local structural content, the other was variation of intensity. By means of either criterion, the author selected sharp parts from input image to form the final image.

Following the G. Hausler and E. Korner idea, R. J. Pieper and A. Korpel established three algorithms to select the sharp part of the input image segments, which then form the final image parts.[5] These three criterions could be called simple average, difference operation, local maximum and local minimum. These methods which we have mentioned above are all implemented without translating the spatial information to the frequency domain. S. A. Sugimoto and Y. Ichioka, following essentially same the idea, proposed a fourth algorithm to select the sharp part of each input image slice.[6] The algorithm was called the local variance weighted averaging method. Since we have implemented the above four algorithms, we will discuss the principles of these four algorithms in chapter four.

Finally, we discuss four new methods to improve the depth of field using digital image filtering in spatial frequency domain. We would like to point out that in 1956, E. L. O'Neill had already found that by using a high pass filter in the spatial frequency

domain, edge sharpening of image was produced. The result can be understood by invoking the analogy between optical filtering and electrical filtering.[7]

Chapter Three Spatial Frequency Domain Methods

3.1 Basic Concept

All methods which are introduced in this chapter are performed within the spatial frequency domain. To transform from spatial domain to spatial frequency domain, we can perform the forward Fourier transform on a image. The inverse Fourier transform takes the image spectrum back to spatial domain. We can use the fast Fourier transform program for these kinds of transformations. In the appendix C, we describe two-dimensional FFT in detail. From now on, we will concentrate on the operations following the forward FFT.

The edges and other sharp transition (such as noise) in the gray level of an image contribute heavily to the high-frequency content of its Fourier transform. This statement can be proved quite easily by experiment. We can filter the spectra over all frequencies. We are able to find out that, only when filtering the high frequency range, the edge information disappear. Otherwise, we can always see the edge of the image, even though we have already suppressed the DC or a low frequency range. Therefore, we

know that after blurring, the spectral energy distribution of a focused image is biased towards lower frequencies range.[8] The blurring can be reduced via the spatial frequency domain by filtering a specified range of low-frequency components in the spectrum of a given image. This is an extremely important point. Actually, it is the starting point of all the methods in this chapter. The other very important point which we would like to point out is that all kinds of digital image processes are, in some sense, image dependent.

All the methods in this chapter can be considered as spatial frequency domain filter methods. The details of the filter in general depend on the specific spectrum characteristics. So that before we design a filter in detail, we must first know the spectrum characteristics of input image set. The most common way which investigates the spectrum characteristics of any image is a histogram.

The general definition of histogram is the plot of $P_r(r_k)$ vs. r_k. Here, $P_r(r_k)$ is the probability of the Kth gray level. The r_k is the Kth gray level of the pixels. In our case, we use the relationship between the spatial frequency bands and the energy spectrum for histogram. Since we represent spatial frequency in terms of number of pixels, the actual relationship is between the number of pixels and the energy spectrum.

In our two-dimensional FFT program, we shift origin of the Fourier spectrum to the center of the image. From the translation properties of the Fourier transform, we have:

$$f(x,y)(-1)^{(x+y)} \leftrightarrow F(u-\frac{N}{2},\ v-\frac{M}{2}) \qquad [3.1]$$

Since the size of our experiment image is 256×128, N equals 128, M equals 256.

Based on the above information, we know that the pixel corresponding to the DC spatial frequency now is located at the pixel which has coordinate (129, 65). Naturally, we set the center of frequency band in pixel (129,65). Since our coordinate system start

from (1, 1) instead of (0, 0), the center is located in (129, 65) not (128, 64). The dimension in the vertical direction is exactly two times longer than in the horizontal direction in our experiment image set. In order to be as accurate as possible, we employed elliptical bands. The two pixels in vertical direction represent same spatial frequency as one pixel in the horizontal direction.

The way to set the spatial frequency band for histogram is based on following elliptical equation:

$$\frac{(I - 129)^2}{4} + \frac{(J - 65)^2}{1} \leq K^2 \qquad [3.2]$$

Here, I is the row; J is the column of image array.

We have already mentioned in the beginning of this chapter that the energy spectrum caused by blurring will mostly concentrate in the low frequency area. Since our purpose is to remove the blur from the image, it is logical that we set more bands in low frequency range than in high frequency range. We find the integer series which is related to Fibonacci's series is suitable for this purpose. The value of K is listed in Table 1 on page 9. Total ten band is set. The number of pixels which has fallen in each band is listed in Table 2 on page 10.

In particular band one and band two, there are seven pixels. In band one (just like we want) only one pixel (129, 65) corresponds to DC spatial frequency. The pixels in band two have coordinates (127,65), (128,65), (129,64), (129,66), (130,65), (131,65). Then we compute the energy spectrum for each band. We will discuss detail in histogram filter method (see section 3.3). Up to now, we have already built up enough background information for the methods in this chapter.

Table 1. The K values for setting the histogram band.

BAND	VALUE OF K
1	0
2	1
3	2
4	4
5	7
6	12
7	20
8	33
9	54
10	bigger than 54

Table 2. The distribution of pixels.

BAND	NUMBER OF PIXELS
1	1
2	6
3	18
4	72
5	206
6	590
7	1616
8	4314
9	11474
10	14471

3.2 DC Suppression

The filter in this method is to make pixel (129,65), which corresponds to the DC term to be zero in each of original semiblured input image. This is the first step. It can be written in following equation:

$$F'(u, v) = F(u, v) - F(0, 0) \qquad [3.3]$$

The second step is that these three input image are added after DC is removed.[7] We can use GIPSY command MERGE to do it. We use expression $(x + y + z)$ for MERGE. More on command MERGE is in appendix 2, GIPSY system.

Since we use following expression for forward FFT

$$F(u, v) = \frac{1}{\sqrt{MN}} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y)\, e^{-j2\pi(\frac{ux}{N} + \frac{vy}{M})} \qquad [3.4]$$

The definition of the average value of a two-dimensional discrete function is:

$$\bar{f}(x, y) = \frac{1}{\sqrt{MN}} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) \qquad [3.5]$$

Substitution of $u = v = 0$ in the above equation yield:

$$F(0,0) = \frac{1}{\sqrt{MN}} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x, y) \qquad [3.6]$$

Therefore, the average $\bar{f}(x, y)$ is related to the Fourier transform of $f(x, y)$ by the equation

$$\bar{f}(x, y) = F(0, 0) \qquad [3.7]$$

So that,

$$F(u, v) - F(0, 0) \leftarrow \rightarrow f(x, y) - \bar{f}(x, y) \qquad [3.8]$$

It means that removing the DC term in spatial frequency domain is equivalent to subtracting the average in spatial domain. For an image having a dominant light background, the average intensity is close to the background level. Just like the theory predicted, the result after these two steps looks like the 'negative' of the desired image. By performing a contrast reversal in the image, we get the final result. For the final result of this method, see Figure 3 on page 25 part e). And for the Fourier amplitude spectrum of this method, see Figure 4 on page 26 part a), b).

3.3 Histogram Filter Method

In histogram method, we first compute the energy spectrum E_k for each spatial frequency band by following formula:

$$E_k = \frac{1}{J_k} \sum_{m} \sum_{n} |F(u, v)|^2 , \qquad k = 1, 2, \ldots, 10 \qquad [3.9]$$

in band k

Where J_k is the total number of pixels in the Kth band. F(u,v) is the Fourier transform of the image slice. We normalize by dividing equation [3.9] by the total energy,

$$R_k = E_k / \sum_{m} \sum_{n} |F(u, v)|^2 , \qquad k = 1, 2, \ldots, 10 \qquad [3.10]$$

all bands

Then we compute the ratio of this quantity H_k between three original input images which have distinct in-focus and out-of-focus regions and the fourth image which is in-focus.

$$H_k = \frac{(R_k)_{focus}}{(R_k)_{blurred}}, \quad k = 1, 2, \ldots, 10 \qquad [3.11]$$

We design different spatial frequency domain filter for each input image based on normalized average energy ratio which is shown in Table 3 on page 14. Any pixel whose coordinate is satisfied by the following equation,

$$K_{(k-1)}^2 < \frac{(I - 129)^2}{4} + \frac{(J - 65)^2}{1} \leq K_k^2 \qquad [3.12]$$

falls into the (k-1)th band.

The histogram method can be explained in three steps. The first step is performing FFT to each input image. The second step is performing individual filter for each input image. The third and also the last step is that of inverse FFT and adding these three images into the final result by using MERGE command. We discuss this GIPSY command in appendix 2.

In order to get the best result, we use expression $(X + Y + Z)$ for MERGE command. See the result in Figure 3 on page 25 part f). And for the Fourier amplitude spectrum of this method, see Figure 4 on page 26 part c), d).

Table 3. The multipliers for the histogram filter.

BAND	RATIO FOR IMAGE 1	RATIO FOR IMAGE 2	RATIO FOR IMAGE3
1	1.0169153	1.0154734	0.9913197
2	0.0247669	0.0280321	0.0392844
3	0.3880279	0.7439507	1.1175056
4	0.5482105	0.6516224	1.0007692
5	0.6968103	0.5772497	0.9440734
6	0.7748620	0.8506678	1.5880096
7	1.0035917	1.140134	2.0797212
8	1.9217687	1.7057594	2.1951584
9	1.6222664	2.2295082	2.6070288
10	0.9608938	1.5221239	2.1772152

3.4 Suspend Band Next to DC

By carefully looking at Table 3 on page 14, we find out that in band two, the ratio[3.11] is much smaller than one for all there images. It indicates that there is significantly more energy in each blurring image as compared to the in-focus image. This extra energy obviously is caused by blurring. The idea to suspend band two of each image is quite natural.

We perform the four steps just like previously methods.(forward FFT, filter, inverse FFT, add images) The only difference is in the second step. Instead of using the histogram filter, we use the band reject filter which suspends band two. Any pixels which have fallen in following range are set to zero.

$$0^2 < \frac{(I - 129)^2}{4} + \frac{(J - 65)^2}{1} \leq 1^2 \qquad [3.13]$$

Of course, we can extend this method by suspending more bands as long as the ratio of these bands shown in Table 3 on page 14 is much small than 1. On the other hand, if we suspend too many bands, we will lose too much information. The result will demonstrate edge detection. We can also suspend DC band or band one and a few low frequency bands. Whenever we suspend DC band, we have to perform a contrast reversal.(see section 3.2) See the result in Figure 3 on page 25 part g). And for the Fourier amplitude spectrum of this method, see Figure 4 on page 26 part e), f).

3.5 Frequency Mixing Method

We first discuss the principle of this method. It is known that multiplying the image intensities by a periodic grating function is equivalent to sampling the image. If blurring is introduced regionally in the image, through either natural or artificial means, the sampling of these regions will disappear. It means that we are only sampling the in-focus part of each input image. Since the blurred parts aren't sampled, these regions do not contribute to the information content in the spatial frequency domain sidebands created through sampling. The in-focus part is sampled, so that in the spatial frequency domain, there will be sideband representations of the signal. The number of sidebands appearing in the frequency domain array will depends on the sampling frequency and the size of the pixel. This follows from equation [*C*.27] and [*C*.28] . We shift the 1th-order sidebands information, which includes only the in-focus contributions to the origin and eliminate the zero-th order signal which contains both in-focus and blur information. We first produce two identical images by applying the grating to an in-focus image which has been obtained by stopping down the camera lens. Then we make, in one of them, the top half part blurred, and in the another, the bottom half part blurred through a convolution process discussed in more detail later. Since the grating effect is destroyed completely in the blurred region, we effectively are only sampling the in-focus part of input image. After applying the horizontal grating lines on the input images, we expect sidebands to appear in the vertical direction on the spectra.

With the aid of FORTRAN mod function, we create the computer grating on the in-focus image. We choose mod two. This means the first row of pixel of this image is not changed. Then make the second row of pixels of this image to be zero. This process is repeated throughout the image. By repeating this method for any two row of pixel,

we get the mod two grating. The reason for us to choose mod two is explained in the following paragraph:

According to the sampling theorem, in order to avoid aliasing, that is, there is no overlap of spectral densities, the sampling frequency has to satisfy the following relationship:

$$f \geq 2B \qquad [3.14]$$

Here, B is the highest frequency of the original signal. There are two more factors we have to consider when we choose the sampling frequency. One is the dimensions of input image in term of pixels. The other is that now we are dealing with the discrete case. In order to get the best results, we have to make the peak value of the adjacent sideband (after sampling) concentrate on single pixel in spatial frequency domain.

Since we have the basic relationship between an ordered sampling of the image and an ordered sampling in the frequency domain of a sampled image.

$$\Delta v = \frac{1}{M\Delta Y} \qquad [3.15]$$

Here, Δv is the sampling interval in the spatial frequency domain,

ΔY is the sampling interval in the spatial domain,

M is number of sampling points in Y direction.

The distance between peak value in each period in terms of the number of rows of the image array can be easily deduced in following way,

$$D = \frac{M}{\text{mod}} \qquad [3.16]$$

Here, D is the distance between two peak values,

M is the size of image in sampling direction, (in our case, it is the Y direction.), and

mod is the distance between two grating lines. All quantity in terms of number of pixels. In our case, M equals 256, mod equals 2.

In order to make the peak value of each period concentrated in a single pixel, we have to choose the grating in terms of number of pixel such that the quotient of M (M equals 256, mod is 2) is an integer. By using the equation[3.16], we know that the distance between two peak values equal to 256/2 = 128. We also know that the DC frequency is located on pixel (129, 65). The sideband peak values are located in column 65 also. They are located in pixel row 1(129-128 = 1) and row 257 (129 + 128 = 257), respectively. Of course, the latter is in the location which is outside of image. We demonstrate this in Figure 1 on page 19. These two incomplete sidebands mainly represent the information contest of in-focus part of the image in the spatial frequency domain. Some part of fundamental period unavoidably overlaps with the sidebands.

We approximately use the half distance between the two peak value in the frequency domain to separate sidebands. Based on the above, we know the zero-th order signal in the spatial frequency domain extends from pixel row 65 ((1 + 129)/2 = 65) to pixel row 193. ((257 + 129)/2 = 193)

The signal in the range from pixel row 1 to pixel row 64 is mainly due to the right part of the in-focus spectrum, see Figure 1 on page 19. The signal in the range from pixels in row 194 to pixels in row 256 is mainly due to the left part of the in-focus spectrum.

By means of convolution, we artificially make the blurring in top and bottom half respectively in each of two input images. See Figure 3 on page 25 part h), i). The definition of discrete convolution of two discrete function f(x) and g(x) is given by,

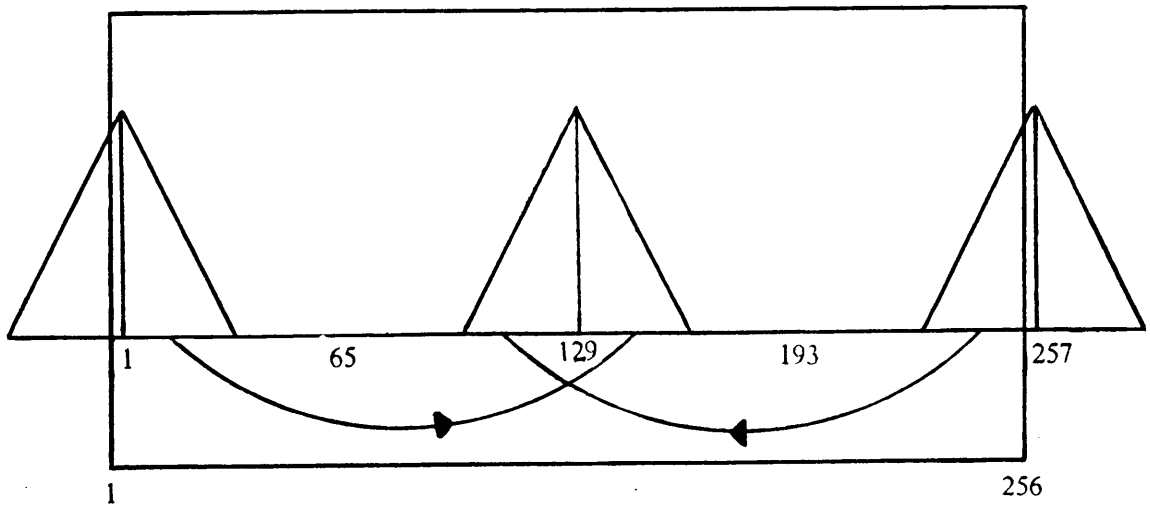


Figure 1. An illustration of intensity shift.

$$f(x) * g(x) \equiv \sum_{m=0}^{M-1} f(m)\, g(x - m) \qquad for \;\; x = 0, 1, 2, \ldots, M - 1. \qquad [3.17]$$

We can easily expand above one-dimensional convolution into a two-dimensional case.

$$f(x,y) * g(x,y) \equiv \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m,n)\, g(x - m, y - n) \qquad \begin{matrix} for \;\; x = 0,1,2, \ldots, M - 1 \\ y = 0,1,2, \ldots, N - 1 \end{matrix} \qquad [3.18]$$

To make enough blurring, we use 11×11 of pixels pyramid. The coefficients in each direction, in sequence, are 0, 0.2, 0.4, 0.6, 0.8, 1, 0.8, 0.6, 0.4, 0.2, 0. See Table 4 on page 21 for details. We convolve these input images with this pyramid to make the one of them, top half blurring, the another bottom half blurring. When we do convolution, we first normalize the all coefficients of this pyramid, then multiply a constant which is a weighted factor for each normalized coefficient. By adjusting weighted factor of the pyramid coefficient to make the background in blurring part as close visually as possible to the background of the unblurred part. The reason for doing this is for the comparision. The final coefficient can be expressed in the following way,

$$C(I,J)_{final} = \left[\frac{C(I,J)}{\sum_{I=0}^{11} \sum_{J=0}^{11} C(I,J)} \right] \times W \qquad [3.19]$$

Where, W is the weighted factor, and the value for C(I, J) can be found in Table 4 on page 21. In our case, W equals 2.1.

The convolution process (blurring) tends to destroy the effect of sampling. The sidebands are essentially correspond of shifted in-focus spectrum. Our algorithm replace the intensity of the pixels in row 129 to row 192 by pixels in row 1 to row 64, and replace intensity of the pixels in row 66 to row 128 by pixel row 194 to 256.

Table 4. The coefficients for convolution.

J \ I	1	2	3	4	5	6	7	8	9	10	11
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.04	0.08	0.12	0.16	0.20	0.16	0.12	0.08	0.04	0.00
3	0.00	0.08	0.16	0.24	0.32	0.40	0.32	0.24	0.16	0.08	0.00
4	0.00	0.12	0.24	0.36	0.48	0.60	0.48	0.36	0.24	0.12	0.00
5	0.00	0.16	0.32	0.48	0.64	0.80	0.64	0.48	0.32	0.16	0.00
6	0.00	0.20	0.40	0.60	0.80	1.00	0.80	0.60	0.40	0.20	0.00
7	0.00	0.16	0.32	0.48	0.64	0.80	0.64	0.48	0.32	0.16	0.00
8	0.00	0.12	0.24	0.36	0.48	0.60	0.48	0.36	0.24	0.12	0.00
9	0.00	0.08	0.16	0.24	0.32	0.40	0.32	0.24	0.16	0.08	0.00
10	0.00	0.04	0.08	0.12	0.16	0.20	0.16	0.12	0.08	0.04	0.00
11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

When we do the row shift, we first multiply the magnitude intensity in pixels from row 1 to 64 and 194 to 256 with a factor two, then shift them to the right place as defined above. The reason for the factor of 2 is because grating is equivalent to the string of pulse or the rectangle function. We have the Fourier transform pair for rectangle function.

$$rect(t/\tau) \leftarrow \rightarrow \tau S_a(\omega\tau/2) \qquad [3.20]$$

The envelope of any signal after sampling in the spatial frequency domain is the S_a (sine over argument) function, that is sin x / x.

In our case, for the pulse function has a period T which is equal to 2 and with duty cycle, τ, equal to 1 as expressed in term of pixels.

After substitution of these values into the above Fourier transform pair, we have,

$$rect(2) \leftarrow \rightarrow S_a(\frac{\omega}{2}) \qquad [3.21]$$

The peak first harmonic is located at ω_0.

$$\omega_0 = \frac{2\pi}{T} \qquad [3.22]$$

So that if the peak value for zero harmonic is 1. then the peak value for first harmonic is by substituting ω_0 into the S_a function.

$$S_a(\frac{\pi}{2}) = \frac{\sin(\frac{\pi}{2})}{\frac{\pi}{2}} = \frac{2}{\pi} = 0.6366197 \qquad [3.23]$$

It means the peak value in the first harmonic is about sixty four percent of the peak value at the origin. So, we should multiply the intensity of each pixel with a factor of

1.57. For reasons that we do not fully understand, a factor 2 worked better. This is continued using an rms calculation as well as a subjective visual test.

The blurring part not only has impact on magnitude, but also has impact on phase. In the process described above, besides shifting magnitude information, we also shift the phase information along with it. The results are extremely good. See Figure 3 on page 25 part j). If we only shift magnitude not phase, the result is not as good as shifting both. By comparing Figure 3 on page 25 part j) with Figure 3 on page 25 part k), we can see this quite obviously. In term of the rms values, it also seems true. The rms values for shifting both magnitude and phase is 11.757, and for shifting only magnitude is 28.798. This clearly shows that our other frequency domain methods are inherently limited, since only the magnitude of the spectra are modified.

Although this method has potential, there is a gap between the demonstration of feasibility and practical application. The primary problem is to generate an optically sampled object with sampling frequency consistent with the digital algorithm. We have to make sure that in terms of number of pixels, the quotient of size of image divided by distance between two grating line is an integer. When we choose to use a physical grating, this turns out to be quite difficult. If we are unable to concentrate the peak in the sideband on a single pixel. The result is unsatisfactory.

If we can find the way to overcome these difficulties, this method definitely will be one of the best methods to improve the depth of focus. we draw a block diagram to illustrate the relationship between the computer experiment and physical experiment in Figure 2 on page 24.

For the final result of this method, see Figure 3 on page 25 part j). And for the Fourier amplitude spectrum of this method, see Figure 4 on page 26 part g), h).

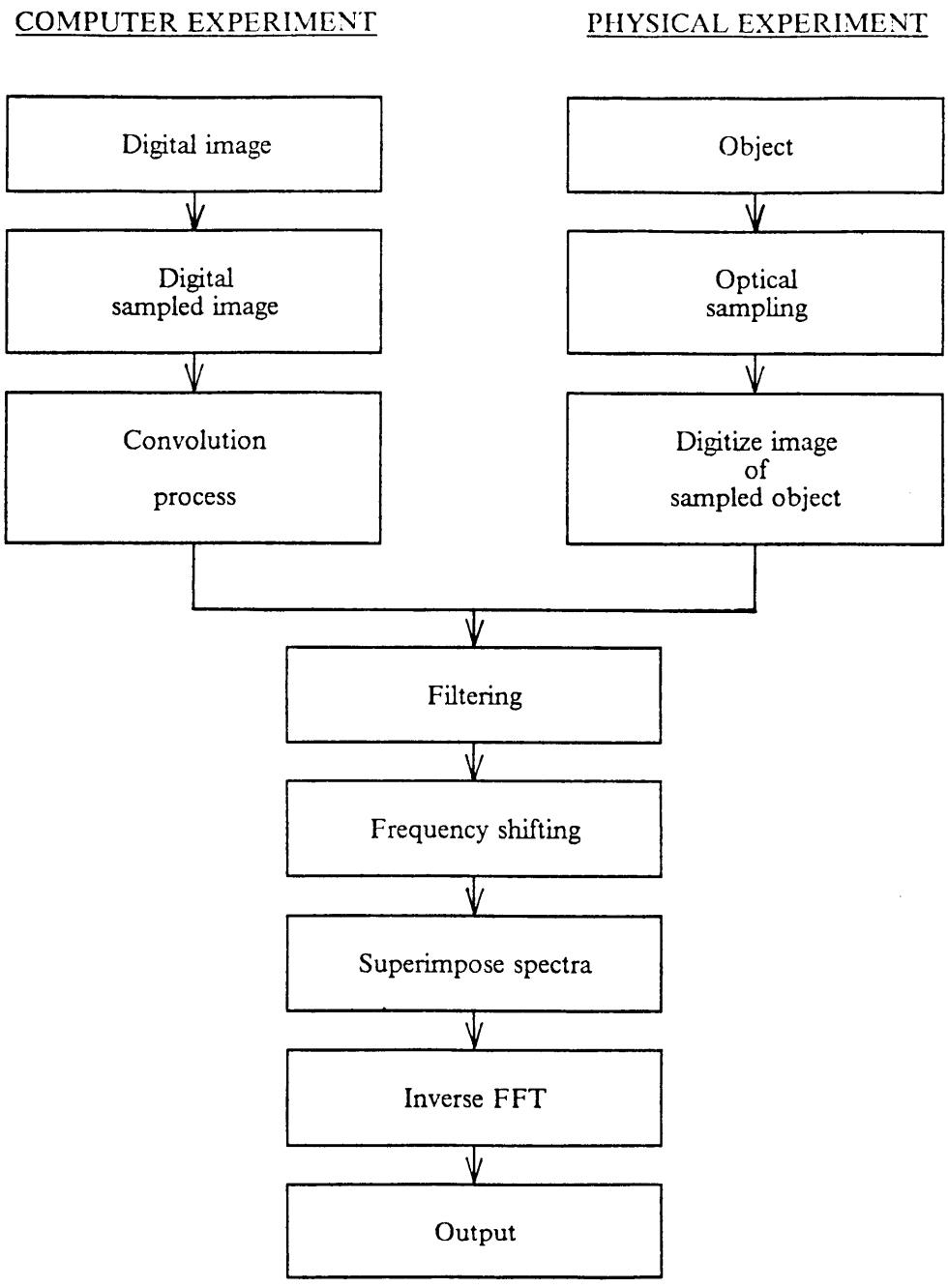


Figure 2. Block diagram of frequency mixing method.

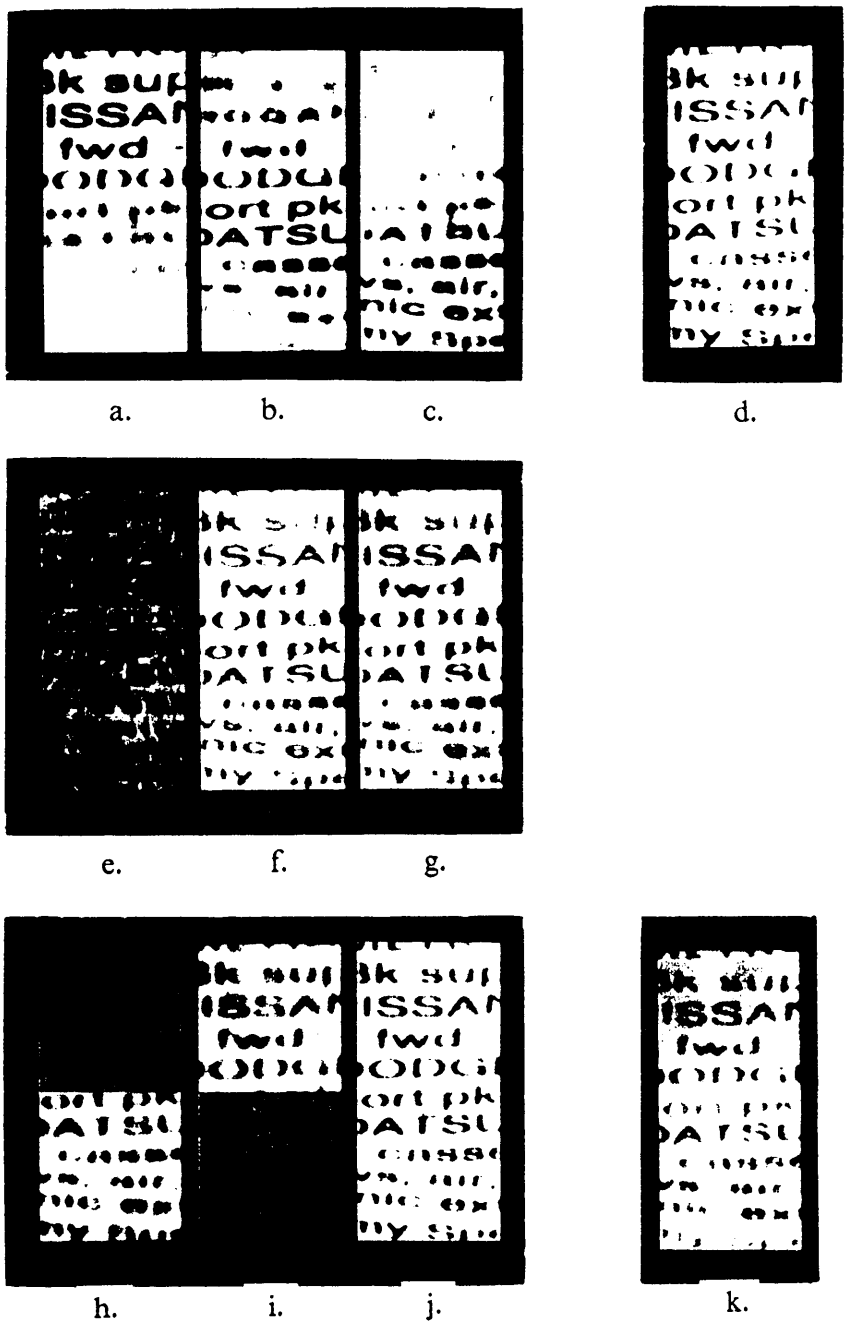


Figure 3. The result four spatial frequency methods.
a), b), c) Three input images. d) All part in-focus image. e) DC suppress. f) Histogram filter. g) Suspend band next to DC. h), i)Input of frequency mixing method. j) Frequency mixing. k) Frequency mixing but no phase shift.

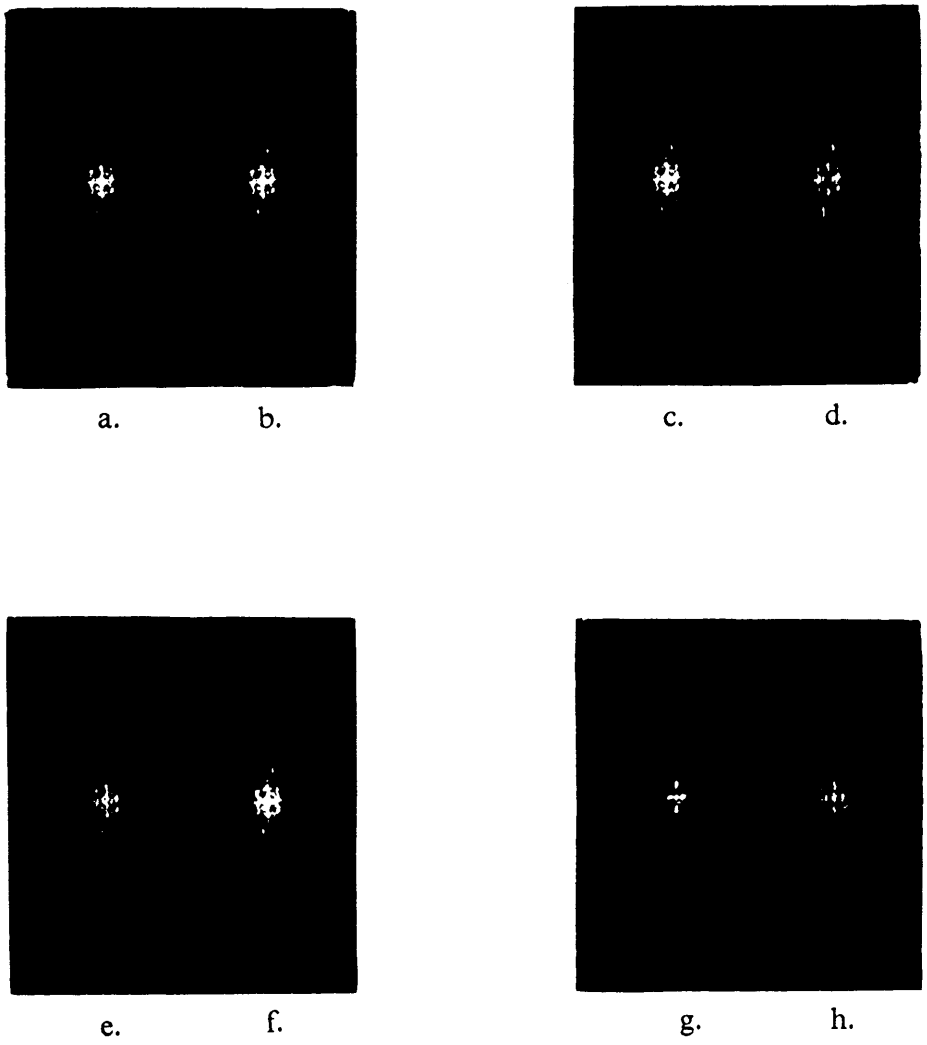


Figure 4. Fourier amplitude spectrum.
for DC Suppress, a)before filtering, b)after filtering. for Histogram Filter Method, c)before filtering, d)after filtering. for Suspend Band Next to DC, e)before filtering, f)after filtering. for Frequency Mixing Method, g)before filtering, h)after filtering.

Chapter Four Spatial Domain Methods

4.1 Basic Concept

In this chapter, we mainly implement four spatial domain methods. The input image set is the same as previously mentioned. See Figure 5 on page 33 part a), b), c). We start with a number of image slices, each having distinct regions that are in-focus and out-of-focus. A composite image is assembled from the in-focus segments of each of these focal-plane depth slices.

All methods which are discussed in this chapter have a common feature, that is, all the methods are performed only in the spatial domain.

4.2 Simple Average Method

This method can be described in following way: First, sum the intensity (gray level) of the pixel which come from different input image slice, but they have the same coor-

dinate pixel by pixel. Then divide by the number of input image slices to obtain the average. Finally, average intensities are used in the final composite image.

Before we formally present this method, we have to establish a mathematical representation for the input image and output image. The remaining three methods which are introduced in this chapter will employ this representation.[5]

P is the number of input image slices.

m runs from 1 to M; M is the horizontal array size in terms of pixels.

n runs from 1 to N; N is the vertical array size in terms of pixels.

p the image slice runs from one to P.

In our case, m equals 256, n equals 128 and P equals 3.

$I_p(m, n)$ is the intensity or gray level of a single pixel which has coordinates (m, n) in pth input image slice. I(m, n) is the intensity of a single pixel which has coordinate (m, n) in the composite image. We use $S_p(m, n)$ to represent the image sampling function which depends on algorithm. Based on these representations, the final composite image pixel I(m, n) satisfies:

$$I(m,n) = \sum_{p=1}^{P} I_p(m,n) S_p(m,n). \qquad [4.1]$$

The algorithm corresponding to this average method therefore can be written in following way,

$$S_p(m,n) = \frac{1}{P} \qquad [4.2]$$

Where after substitution of [4.2] into [4.1] , it follows that:

$$I(m,n) = \sum_{p=1}^{3} \frac{I_p(m,n)}{P} \equiv \bar{I}_p(m,n) \qquad [4.3]$$

For the input set in Figure 5 on page 33 part a), b), c), the result are shown in Figure 5 on page 33 part e).

4.3 Maximum and Minimum Method

This method [5] is based on the following intuitive concept. In order to assemble the final image, we must detect condition of local focus in each slice. By intuition, we can roughly make an approximation that when a pixel is in focus, it exhibits either a maximum or a minimum in the intensity. That means as focus is approached, pixel intensity will either wax or wane monotonically. Our choice of pixel intensity in the composite image is based on which pixel has the biggest absolute difference between its pixel intensity and the average intensity of the pixels for coordinate (m, n). We can represent this algorithm with the in following formulas:

$$S_p(m,n) = 1 \qquad for\ p = p_0(m,n) \tag{4.4a}$$

$$S_p(m;n) = 0 \qquad for\ p \neq p_0(m,n) \tag{4.4b}$$

Here, $p_0(m, n)$ is the pixel which has coordinate (m,n) and also is closest to being in focus.

$$I(m,n) = \sum_{p=1}^{3} I_p(m,n)S_p(m,n). \tag{4.5}$$

For the input set in Figure 5 on page 33 part a), b), c), the result are shown in Figure 5 on page 33 part f).

4.4 Difference Operator Method

This method uses the information taken from neighboring pixels in the same P_{th} image. It is based on the concept that lines are sharper when in focus. This corresponds to preservation of the high frequency terms. From linear transform theory, differentiation in spatial domain is equivalent to multiplication by the frequency in the frequency domain.

$$\frac{dh(t)}{dt} \leftarrow \rightarrow j\pi f H(f) \qquad [4.6]$$

So that the differential operation weights the higher frequencies more than lower frequencies. In a discrete case, we have to use a difference operation instead of differential operator in a continuous case. The definition of a nondirectional Sobel operator[8] $D_p(m, n)$ is the following[4.7]:

$$\begin{aligned} D_p(m,n) = &\,|I_p(m-1,n+1) - I_p(m+1,n-1)| \\ &+ |I_p(m+1,n+1) - I_p(m-1,n-1)| \\ &+ |I_p(m,n+1) - I_p(m,n-1)| \\ &+ |I_p(m-1,n) - I_p(m+1,n)| \end{aligned} \qquad [4.7]$$

The algorithm then can be formed in following way,

$$D_{p_0} = \max\{D_1(m,n), D_2(m,n), D_3(m,n)\} \qquad [4.8]$$

$$S_p(m,n) = 1 \qquad for\ p = p_0(m,n) \qquad [4.9a]$$

$$S_p(m,n) = 0 \qquad for\ p \neq p_0(m,n) \qquad [4.9b]$$

and,

$$I(m,n) = \sum_{p=1}^{3} I_p(m,n)S_p(m,n) \qquad [4.10]$$

We interpret the above formula into words is that, firstly, we have to decide the size of local area. Here, the local area means the size of neighboring pixels. Secondly, after performing the difference operation pixel by pixel on each input image slice, the maximum difference operator is computed. Finally, the pixel which satisfies [4.8] is placed into the resulting image. In order to solve the problem relating to the edge of the image, we extend the image size from 256×128 to 258×130, and set to all the extra pixels to zero intensity. For the input set in Figure 5 on page 33 part a), b), c), the result are shown in Figure 5 on page 33 part g).

4.5 Local Variance Weighted Averaging Method

The method [6] also uses the information about neighboring pixels much like the previous method. When the amount of defocusing of a image increases, the high spatial frequency terms will decrease remarkably. To convert the above concept into a spatial domain operation, we have the following: if the image is in-focus, the local variation in the image is high. When the image is out-focus, the local variation in the image is low. We select the pixels at coordinate(m, n), which has a maximum in the local variance over the P image slices. These are then put into the output image in the corresponding pixel coordinates (m, n).

We choose the three by three square area for the local variance method and extend each input image two row and two columns in exact same way with the previous method. The algorithm of this method then can be written in following expression:

$$\bar{I}_p(m,n) = \frac{1}{9} \sum_{p=-1}^{1} \sum_{q=-1}^{1} I_p(m+p, n+q) \qquad [4.11]$$

Where, $I_p(m,n)$ is the local average.

$$V_p(m,n) = \frac{1}{9} \sum_{p=-1}^{1} \sum_{q=-1}^{1} [I_p(m+p, n+q) - \bar{I}_p(m,n)]^2 \qquad [4.12]$$

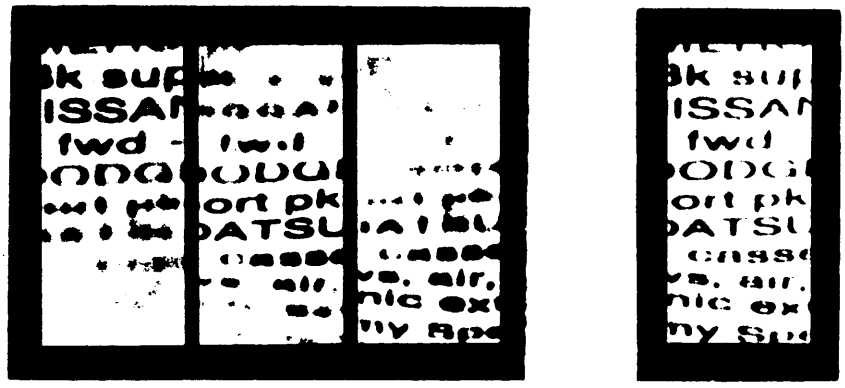
Where, $V_p(m,n)$ is local variance.

$$W_p(m,n) = \frac{V_p(m,n)}{\sum_{p=1}^{3} V_p(m,n)} \qquad [4.13]$$

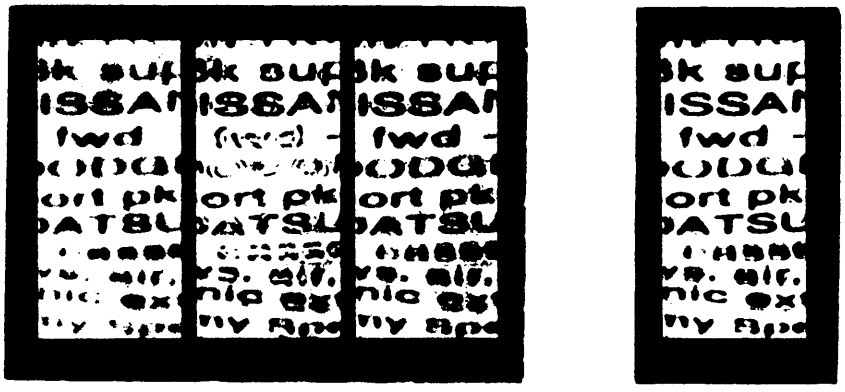
Where, $W_p(m,n)$ is weighted factor.

$$I_p(m,n) = \sum_{p=1}^{3} W_p(m,n) I_p(m,n) \qquad [4.14]$$

From the above formula, we know that the first step is to compute the local variance for each pixel at the same coordinate but different slice by using equation [4.12]. The second step is to compute the weight factor by using equation[4.13]. The last step is to evaluate the pixel intensity according to equation[4.14]. The results for input set Figure 5 on page 33 part a), b), c), are shown in Figure 5 on page 33 part h).



a. b. c. d.



e. f. g. h.

Figure 5. The result of four spatial methods.
a), b), c) Three input images. d) All part in-focus image e) Simple average. f) Maximum and minimum. g) Difference operator. h) Local variance weighted average.

Chapter Five The Comparision of Different Methods

In the previous two chapters, we have implemented a total of eight methods. In this chapter, we try do a comparision. In the beginning of this chapter, we would like to emphasize one very important point. We can make a qualitative evaluate of an image. we will say this image is good or not good or even bad. But this judgment may not consistent with the other more mathematical criterion. The comparisions in this chapter refers strictly to the mathematical concept but not to the human visual system. Sometimes, these two kind of comparisions are consistent with each other, sometimes, they are not. In fact, at present, our lack of knowledge about visual perception precludes the general agreement of mathematical models to visual interpretation.[8] In this chapter, we discuss comparision based only on mathematical quantitative criterion.

5.1 RMS Values

We use RMS value for a fidelity criterion to measure the quantity of the reconstructed image. The RMS is the abbreviation of square-root-of mean square. So that, the fidelity criterion is the root-mean-square error between the input image and output image.

The error between a reference pixel with intensity R(x, y) and a final composite pixel with intensity C(x, y):

$$e(x,y) = R(x,y) - C(x,y) \qquad [5.1]$$

The squared error averaged over the image array is:

$$\bar{e}^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} e^2(x,y) \qquad [5.2a]$$

That is,

$$\bar{e}^2 = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [R(x,y) - C(x,y)]^2 \qquad [5.2b]$$

The RMS error is:

$$e_{rms} = \sqrt{\bar{e}^2} \qquad [5.3]$$

We use GIPSY command RMS to compute rms error between final composite image and the image which we take for the comparision purpose.

We use the Virginia Tech newspaper 'Collegiate Times' for the object of the input image set. See Figure 5 on page 33 part a), b), c). This set has three image. Each of them has a distinct range that in-focus and out-of-focus. Each input of this set con-

tributes approximately 33% to in-focus part of the final composite image. For comparision purposes, we take the in-focus reference image from same object by stopping down the camera lens and and turning up the lights. See Figure 5 on page 33 part d). All the images in the experiment have 256 row pixels and 128 column pixels.

5.2 Results

We summarize the results in Table 5 on page 37. From this table, we know that the best method with respect to the RMS value is frequency mixing method. The second place belongs to simple average method in spatial domain.

The rank of methods depends on the characteristics of the original image set, as well as the filtering. We should say that this is just a first step in this area. The result of all the eight method are presented in Figure 3 on page 25 in chapter three and in Figure 5 on page 33 in chapter four.

Table 5. A comparision using RMS values.

Spatial Method		Spatial Frequency Method	
Method	RMS Value	Method	RMS Value
Simple Average Method	32.578	DC Suppression	89.545
Maximum and Minimum Method	57.485	Suspend Band Two	31.523
Difference Operator Method	53.802	Histogram Filter Method	34.010
Local Variance Weighted Average Method	48.723	Frequency Mixing Method	11.757

Chapter Six Conclusion and Summary

In this thesis, we discuss four types of new methods to extend the depth of field by using digital image filtering. They all show various degrees of improvement. Among them, the result of frequency mixing is very encouraging. Of course, there is a gap between the feasibility and practical application as a by-product of the frequency mixing approach.

We found that the energy shift caused by blurring appears in the low frequencies. It is mainly concentrated in the frequency range next to DC. This is the bases for the suspend band next to DC algorithm discussed in chapter three.

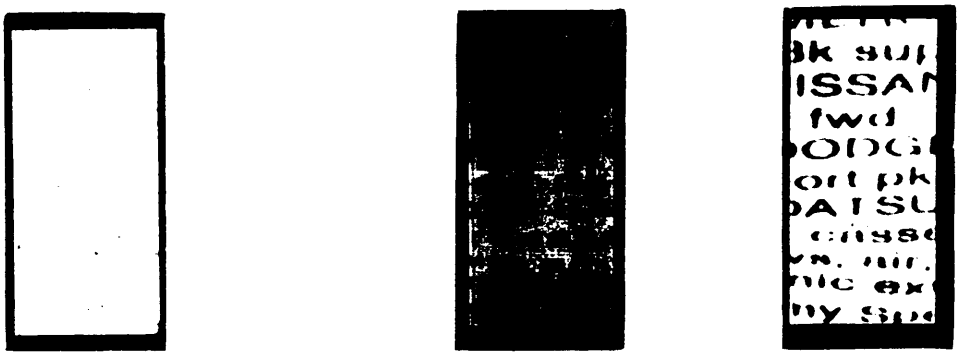
We found that the blurring effects not only magnitude of the image spectrum, but also the phase information. See section 3.5.

We have implemented four spatial methods based on previous work in chapter four. The common characteristics of these eight methods, which is discussed four of them in chapter three and four of them in chapter four, is that the final result is the composite of a set of images each of which has in-focus as well as out-of-focus regions. A comparision among them has been made. This is the first step in this area. It is quite important, because we try to evaluate the different methods which improve the depth

of field. In this area, it is very important to find a criterion for evaluating the image which is at least in some degree compatible for both mathematical and human visual systems. Right now, as we pointed out in this chapter, the RMS value is not always consist with the subjective interpretation.

Since the FFT has a variety applications in many areas, this thesis includes a discussion of the 2DFFT algorithm as applied to rectangle arrays.

Finally, we would like to point out several interesting phenomenon and problems for further investigation. The first is that if we suspend enough bands except DC band in the spatial frequency domain, the resulting image will shown edge detection. See Figure 6 on page 40 part a). This may be useful for edge enhancement. The second and third problems are related to the frequency mixing method. As we mentioned in section 3.5, for reason which we do not fully understand, the factor 2 works better than the factor 1.57 which was predicted by the theory. This can be seen by comparing Figure 6 on page 40 part b) which corresponds to the factor 1.57 and part c) which corresponds to the factor 2. The third problem is the development of a methodology for application of a physical grating. We have already pointed out in chapter three, when using the physical grating, the result was quite poor. For comparison, the picture of the input image after putting a physical grating on is shown in Figure 6 on page 40 part d). The resulting image of the frequency mixing method is presented in Figure 6 on page 40 part e).



a. b. c.

d. e.

Figure 6. The phenomenon and problems.
a) edge detect effect. b) image for factor 1.57. c) image for factor 2 d) input image after putting physical grating. e) the result image for physical grating.

Bibliography

[1]M. Mino and Y. Okono, "Improvement in the OTF of a Defocused Optical System Through the Use of Shaded Apertures", Appl. Optics, Vol. 10, no. 10, pp. 2219-2225, Oct. 1971.

[2]J. T. McCrickerd, "Coherent Processing and Depth of Focus of Annular Aperture Imaginery", Appl. Opt., Vol. 10, no. 10, pp. 2226-2230, Oct. 1971.

[3]G. Hausler and E. Korner, "Expansion of Depth of Focus by 'Image Depuzzling'", Proceedings of 6th International Conference on Pattern Recognition, Oct. 1982.

[4]G. Hausler, "A Method to Increase the Depth of Focus by Two Step Image Processing", Opt. Comm., Vol. 6, no. 1 pp. 38-42, Sep. 1972.

[5]R. Pieper and A. Korpel, "Image Processing for Extended Depth of Field", Appl. Optics, Vol. 22, no. 10, pp. 1449-1453, May 1983.

[6]S. A. Sugimoto and Y. Ichiota, "Digital Composition of Images with Increased Depth of Focus Considering Depth Information", Appl. Optics, Vol. 24, no. 14, pp. 2076-2080, July 1985.

[7]E. L. O'Neill, "Spatial Filtering in Optics", IRE Trans. on Information Theory, pp. 56-65, June 1956.

[8]P. C. Gonzalez and P. Wintz, DIGITAL IMAGE PROCESSING, Addison-Wesiey Co., 1977.

[9]A. Rosenfeld and A. C. Kak, DIGITAL PICTURE PROCESSING, Academic Press, Inc, 1976.

[10]E. O. Brigham, THE FAST FOURIER TRANSFORM, Prentice-Hall, Inc., 1974.

[11]F. G. Stremler, INTRODUCTION TO COMMUNICATION SYSTEM, Addison-Wesley Co., 1982.

[12]J. W. Goodman, INTRODUCTION TO FOURIER OPTICS, McGraw-Hill Co., 1968.

[13]E. Garland and R. W. Ehrich, A GIPSY PRIMER, VPI. 1986.

Appendix A. The Depth of Focus and Resolution in Optical System

The purpose of this appendix is to explain the following statement which is used in chapter one. The wider the aperture of an imaging system, the lower is the depth of field and the higher the resolution. To avoid involved mathematics, we can explain the above statement through physical considerations. According to the ray theory, the wider aperture of an imaging system, the lower is the depth of field. On the other hand, according to wave theory, the resolution increases with aperture width. From geometric optics, we have very basic relationship for a simple imaging system:

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f} \qquad [A.1]$$

Here, a is the distance between the object and lens.

b is the distance between the image and lens.

f is the focus of the lens.

From the Figure 7 on page 45, we examine the rays from a single point of the object. In the focal plane, we can expect to get a point image. When we move the plane around the focus, we can see that the image is no longer a single point. Distortion has occurred. This phenomenon is the so-called circle of confusion. The farther from the focus, the worse the image.

Under the same condition, by moving the plane from focus the same distance, a wider aperture lens produces a more serious distortion. This is quite obvious only by comparing the ray diagrams in Figure 8 on page 46.

The fact that wider aperture of an imaging system leads to a higher resolution is a consequence of diffraction theory. If we put a plane wave in front of a lens, we are able to get a single point, in focus, on the another side of lens. Actually, this spot is not an ideal point. It has width. The width of spot in reverse proportion to the aperture of lens[13]. A standard analytic treatment for the incoherent impulse response report can be found in most texts on optical information processing.[13] The depth of focus in most camera systems is essentially determined by geometrical optics not diffraction optics.

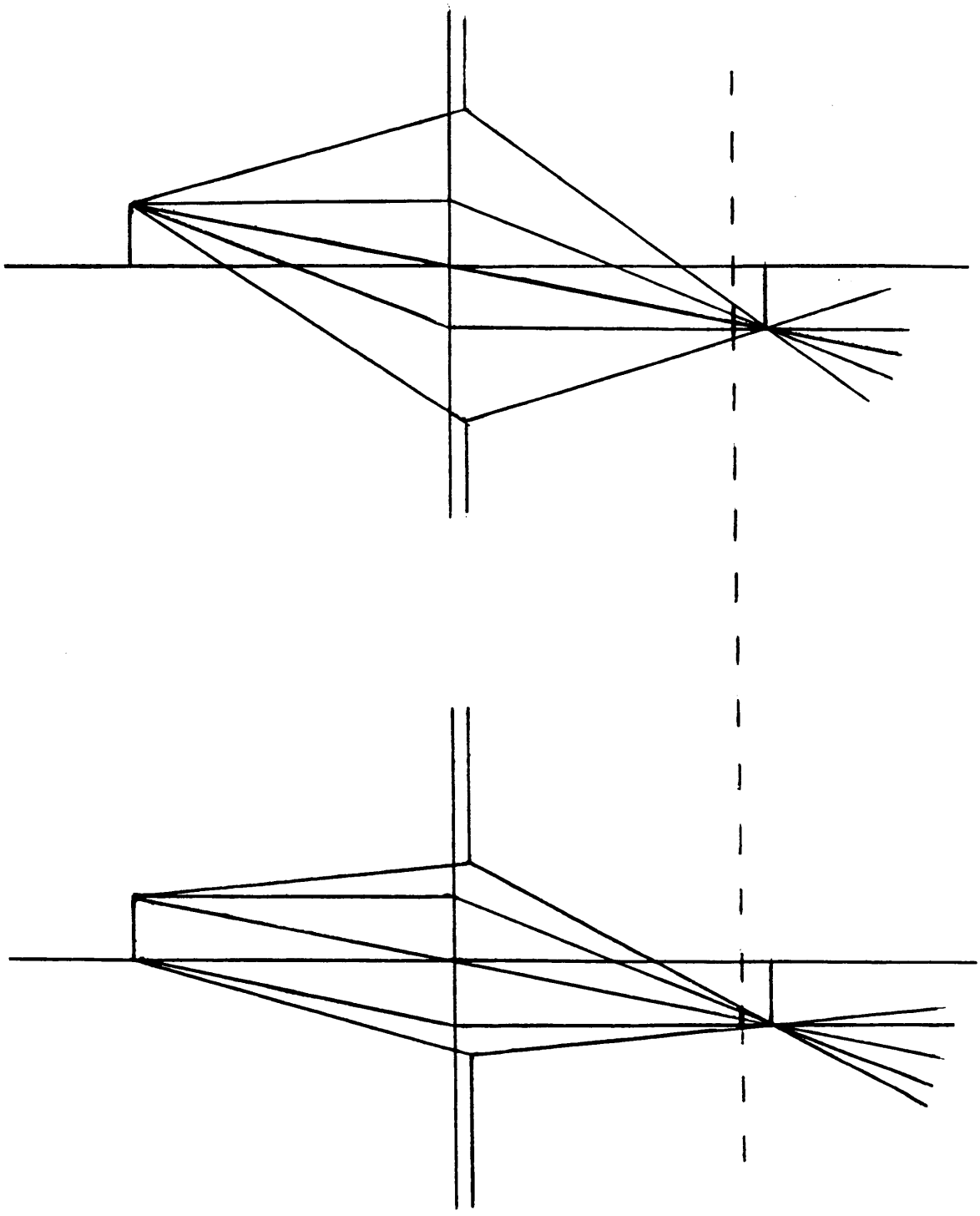


Figure 7. The circle of confusion.
The aperture in top figure is bigger than in lower figure. The circle of confusion in top figure is also bigger than in lower figure.

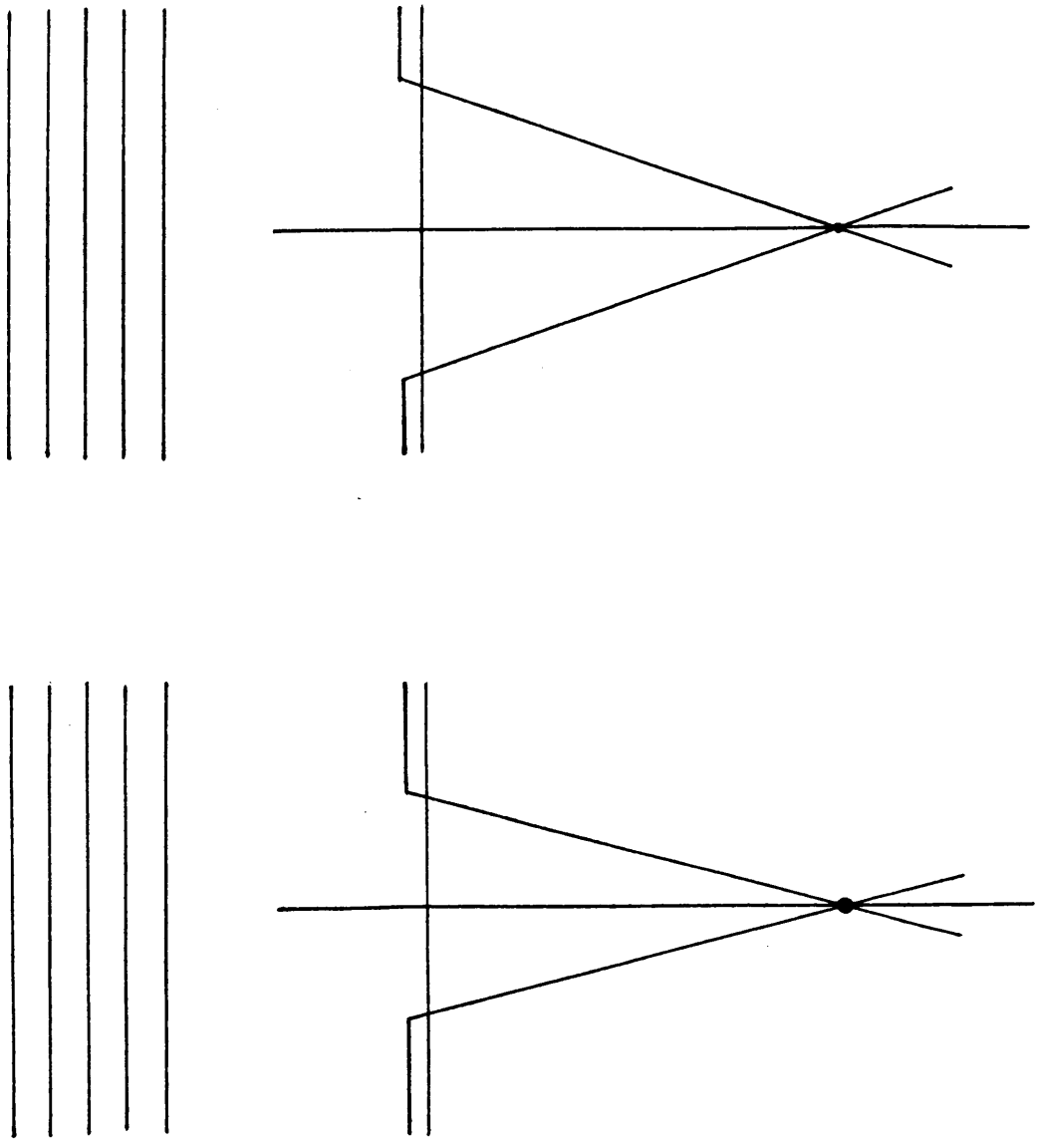


Figure 8. The diffraction theory.
The aperture in top figure is bigger than in lower figure, but the width of spot in top figure is smaller than in lower figure.

Appendix B. GIPSY System

GIPSY [13] is an interactive image processing software package. Each GIPSY command is a collection of subroutines written in RATFOR which is an extension to FORTRAN. With the aid of GIPSY system, we did the digital image experiment.

The hardware part of this image processing system is represented in Figure 9 on page 48. The computer is Digital's VAX 11/780. The computer controls the processes between all part of the system and provides the memory space for storing data (images) and other files (programs). Of course, the VAX computer also handles computation. We put the object which we want to take the image of under the VIDICON camera. The VIDICON and VICOM monitor work together to capture, digitize and store images in the standard image format (SIF). The SIF is designed so that all information about an image is contained in the image file itself. In order to gain access to it, we do not need to know anything about it. By means of a television monitor, we are able to display any image. Using the GIPSY system, we can take any picture of an image which is displayed on the television monitor. See Table 6 on page 51.

A digital image is a quantified representation of an image mapped into a finite two-dimensional array of resolution cells which are called the pixels. The maximum ar-

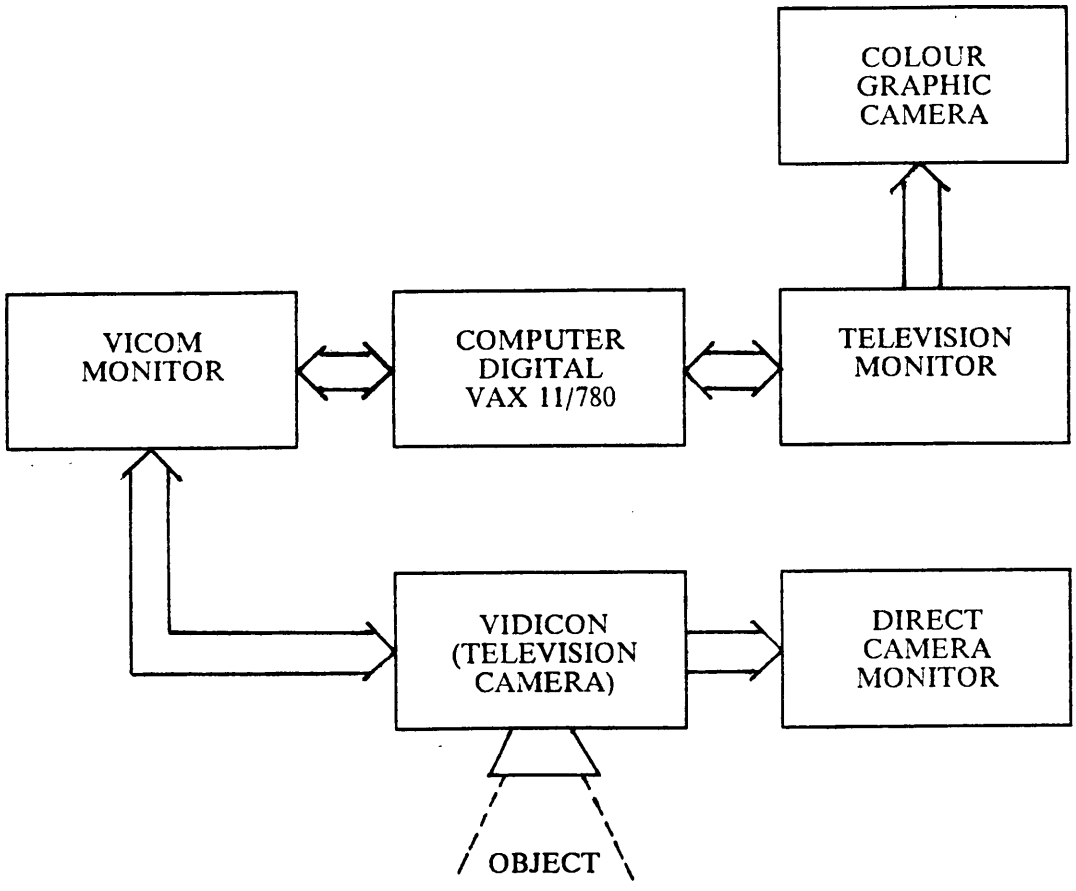


Figure 9. Digital Image Processing System (GIPSY).

ray size is 512×512. Associated with each pixel is a representative gray level intensity. The input and output image is expressed in integer mode. The gray level intensity range is from 255 (light) to 0 (dark).

We use FORTRAN programs to do some important image operations. The GIPSY commands SIFCHR, CHRSIF are used to transform between the ordinary data file and the standard image format(SIF). The function of SIFCHR is converting SIF image into a character file. The function of CHRSIF is converting a character file to a SIF file. We use EXSIF command to make sure that for any image in our experiment, the range of gray shade intensity is from 0 to 255.

Another very important GIPSY command is the MERGE command. The function of MERGE command is that merge up to three SIF files using a user entered function. When using this command, you will be asked the maximum and minimum intensity values for output image. Any values outside this range are clipped to the minimum and maximum values respectively.

Provided below is the interactive sense using GIPSY [13] commands for capturing and displaying on the VICOM monitor a digitized image. The interested reader should review the GIPSY general manuals [13] for other operations.

After we log on the VAX computer, we are in the system level. In order to use GIPSY software package, we have to first get into GIPSY level by entering the command GIPSY. The sign 'G'in the screen indicates that we are now in GIPSY level. There are three main steps. The first step is entering GIPSY command GETDSP VICOM. The function of this command is simple to allocate VICOM device to you. Next step is entering GIPSY command VICOM. The VICOM is a command to handle interactive image processing on the VICOM device. This is the command which we use for capturing, digitizing and storing the image. When we enter this command, we are able to see the *> sign. From now on, we have to enter some subcommands. They are

INT, CAM 1, DIS 1, DIG 1, WIM(filename.SIF), END, in order. Finally, we get back to GIPSY level again. Now, we have already put the resulting image under the filename which we pick up whatever we want. In general, we use SIF for filetype. We have to enter command RELDSP VICOM to release the VICOM device. If we want to take a look at the image which we capture on any display device, we have to enter GETDSP LIGHT(n), n is the code which represents particular display device. Then enter DSPLAY fn.SIF > LIGHT, we should able to see the image. Of course, after this is done, we also need to release the display device. We can use these individual commands to program a runfile for convenience.

The process in which to log on computer by capturing an image and to logoff from the computer is in the following table step by step.

Table 6. The process for capturing an image.

INTERACTIVE SESSION

COMPUTER	USER RESPOSE
	PRESS ENTER KEY
USERNAME:	YOUR USERID(enter)
PASSWORD:	YOUR PASSWARD(enter)
TERMINAL:	VT100(enter)
$:	GIPSY(enter)
G:	GETDSP VICOM(enter)
G:	VICOM(enter)
*>	INT(enter)
*>	CAM 1(enter)
*>	DIS 1(enter)
*>	DIG 1(enter)
*>	WIM 1(fn.SIF)(enter)
*>	END(enter)
G:	RELDSP VICOM(enter)
G:	EXIT(enter)
$:	LOGOFF(enter)

Appendix C. Two-dimensional FFT

For reference, we include a discussion of the FFT algorithm which was written to carry out the operation in this thesis.

The fast Fourier transform is an algorithm for computation of the Fourier transform. This algorithm can compute the discrete Fourier transform more rapidly than any other available algorithms. In 1965, Cooley and Tukey presented this algorithm in the paper "An Algorithm for the Machine Calculation of Complex Fourier Series."[10]

For the maximum efficiency of computation, the number of sample points to be integer power of two, that is, the sample points $N = 2^L$. The ratio of direct Fourier transform to FFT computing time is given by[10]

$$\frac{2N}{\log_2 N} \qquad [C.1]$$

From above formula, we can easily see that the more sampling points we have, the more efficient the FFT becomes.

The one-dimensional discrete Fourier transform can be defined as

$$F(u) \equiv \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x) e^{-j2\pi \frac{xu}{N}} \qquad u = 0,1,2, \ldots N-1 \qquad [C.2]$$

Where u indexes the transformed data, F(u).

N is the number of data points.

x indexes the data set f(x).

Similarly, the data set f(x) can be obtained from the set F(u) using the inverse one-dimensional discrete Fourier transform that can be defined as:

$$f(x) \equiv \frac{1}{\sqrt{N}} \sum_{u=0}^{N-1} F(u) e^{+j2\pi \frac{xu}{N}} \qquad x = 0,1,2, \ldots N-1 \qquad [C.3]$$

The following weight function can be defined,

$$W_N^{xu} \equiv e^{-2\pi j \frac{xu}{N}} \qquad [C.4]$$

These functions satisfy the orthogonal condition,

$$\frac{1}{N} \sum_{x=0}^{N-1} W_N^{xu} W_N^{xu'} \equiv \delta_{uu'} \qquad [C.5]$$

Equation [C.5] can be proven using the representation for a finite geometric series. Consistency of transform [C.2] and [C.3] can be easily shown by substituting [C.3] into [C.2].

The application of the DFT to actual data in the space domain motivates the following definition.

$$\Delta x \equiv \frac{L_x}{N} \qquad [C.6]$$

Where L_x is the length sampled and Δx is the sampling interval in the space domain.(pixel) It can be shown [8] that the sample interval in the spatial frequency domain, Δu, satisfies

$$\Delta u = \frac{1}{N\Delta x} = \frac{1}{L_x} \qquad [C.7]$$

These relation are geometrically interpreted in Figure 10 on page 55.

The translation property of the Fourier transform can be applied to shift the DC point of the frequency domain (1,1) to the center pixel (65,129). The following are transform pairs:

$$f(x,y)(-1)^{x+y} \leftrightarrow F(u - \frac{N}{2}, v - \frac{M}{2}) \qquad [C.8]$$

Where N = 128 and M = 256.

A two-dimensional rectangle DFT can be defined in term of spatial indices x and y as well as the spatial frequency indices u and v as:

$$F(u,v) \equiv \frac{1}{\sqrt{NM}} \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} f(x,y) e^{-j2\pi(\frac{xu}{N} + \frac{yv}{M})} \qquad [C.9]$$

Where N and M are the number of samples taken in the X and Y direction respectively. After minor adjusting in transform [C.9], we see that the 2-D DFT [C.9] can be thought of as two segmental 1-D discrete transform.

$$F(u,v) = DFT_x\{DFT_y f(x,y)\} \qquad [C.10]$$

It follows that the 2-D fast Fourier transform can be implemented as a sequence of two 1-D fast Fourier transform. For this reason, we can concentrate on the one-dimensional aspects of this problem.

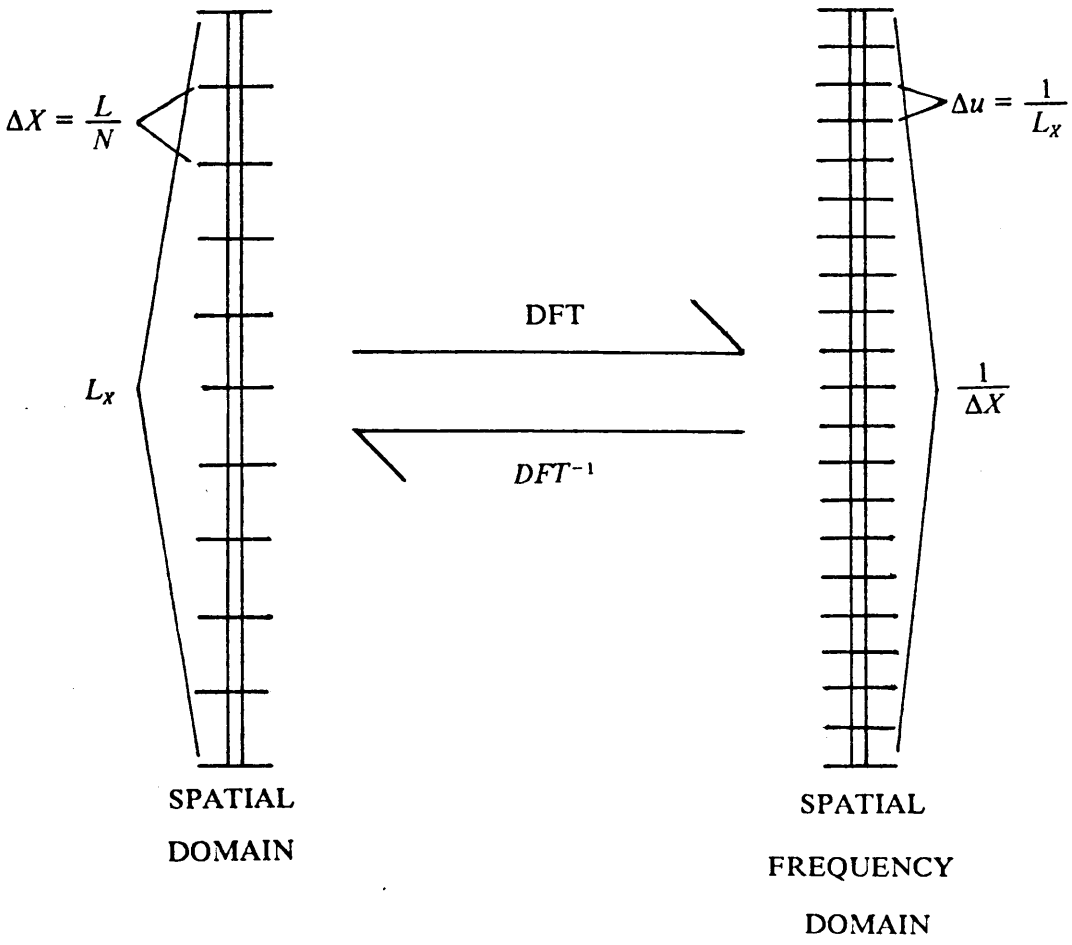


Figure 10. Geometric interpretation of 1-D transform parameter.

Staring from [C.2] and [C.4], it follows that:

$$F(u) = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} f(x) W_N^{xu} \qquad [C.11]$$

This operation can be broken into two subtransforms separating the even and odd sequence of f(x):

$$F(u,v) = \frac{1}{\sqrt{N}} \sum_{x=0}^{\frac{N-1}{2}} (f(2x) W_N^{2xu} + f(2x+1) W_N^{(2x+1)u}) \qquad [C.12]$$

Using the definition [C.4] for the weight function, it is easy to show:

$$W_N^{2xu} = W_{N/2}^{xu} \qquad [C.13a]$$

$$W_N^{(2x+1)u} = W_N^u W_{N/2}^{xu} \qquad [C.13b]$$

For later reference we can define:

$$K = \frac{N}{2} \qquad [C.14a]$$

$$\{f_e: s = 0, \ldots K-1\} = \{f(2x): x = 0,1,2, \ldots K-1\} \qquad [C.14b]$$

$$\{f_o: s = 0, \ldots K-1\} = \{f(2x+1): x = 0,1,2 \ldots K-1\} \qquad [C.14c]$$

$$F_e(u) \equiv \frac{1}{\sqrt{K}} \sum_{x=0}^{k-1} f_e(x) W_k^{xu} \qquad [C.14d]$$

$$F_o(u) \equiv \frac{1}{\sqrt{K}} \sum_{x=0}^{k-1} f_e(x) W_k^{xu} \qquad [C.14e]$$

Equation [C.12] can be written as

$$F(u,v) = \frac{1}{\sqrt{2}}(F_e(u) + W_N^u F_o(u)) \qquad [C.15]$$

The process as discussed is extendable to each term $F_e(u)$ and $F_o(u)$ as shown below:

$$F_e(u) = \frac{1}{\sqrt{2}}(F_{ee} + W_{N/2}^u F_{eo}) \qquad [C.16a]$$

$$F_o(u) = \frac{1}{\sqrt{2}}(F_{oe} + W_{N/2}^u F_{oo}) \qquad [C.16b]$$

When, for example, F_{ee} is the DFT computed from the even term in the set$\{f_e\}$ defined in [C.14b]. The subscript 'o' indicates the odd term. At each splitting the number of term in the subtransforms is cut in half. This can continue until $F(u)$ is written as a linear combination of N, one term, transform. Note, from [C.2], if N = 1, F(0) = f(0).

For illustration, this process appears below for N = 4:

$$F(u) = \frac{1}{\sqrt{2}}\left[\frac{1}{\sqrt{2}}(F_{ee} + W_{\frac{N}{2}}^u F_{eo}) + \frac{1}{\sqrt{2}} W_N^u (F_{oe} + W_{\frac{N}{2}}^u F_{oo})\right], \; u = 1, 2, 3, 4 \qquad [C.17]$$

Where for n = 4 it follows that:

$$F_{ee} = f(0) \qquad [C.18a]$$

$$F_{eo} = f(2) \qquad [C.18b]$$

$$F_{oe} = f(1) \qquad [C.18c]$$

$$F_{oo} = f(3) \qquad [C.18d]$$

Barring the $\frac{1}{\sqrt{2}}$, a graphical representation (known as a butterfly diagram) of the algebraic process [C.17] is shown in Figure 11 on page 59. The scrambled sequence is obtained by reversing the binary decimal representation. Level one represents the operation [C.16], while level two represents operation [C.15]. There will be L level as previous defined, $L = \log_2 N$.

At level one use was made of the fact than

$$W_{N/2}^{(u-N/2)} = (W_{N/2}^{u})\, W_{N/2}^{(-N/2)} = W_{N/2}^{u} \qquad [C.19]$$

The solid lines are the unweighted terms. While the dashed line require weights as indicated on the Figure 11 on page 59. For example, from the Figure 11 on page 59, it follows that

$$F(3) = (\frac{1}{\sqrt{2}})^2 [(X(0) + W_2^1 X(2)) + W_4^3 (X(1) + W_2^1 X(3))] \qquad [C.20]$$

Patterns in N = 4 case which are generally valid are mentioned below:

1. The number of level L:

$$L = \log_2 N \qquad [C.21]$$

2. Each level is separated into divisions. The number of division per level 'D' is given by:

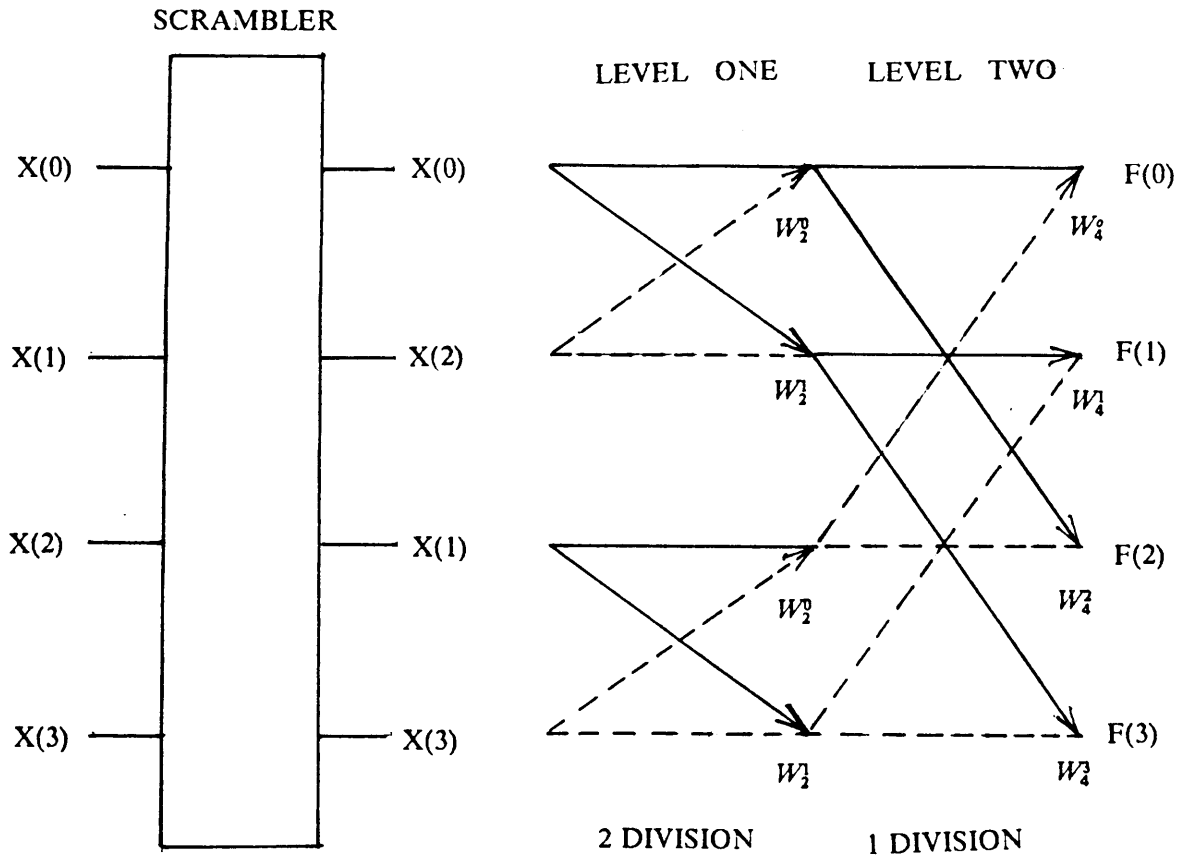


Figure 11. Butterfly diagram for N = 4.

$$D = \frac{N}{2^l} \qquad [C.22]$$

Where l is the level number $l = 0,1,2, \ldots L$.

3. Each term is determined by the linear combination of two term in the previous level. The dual term appears in the same division and is always from the same two term. For example, on Figure 11 on page 59, F(0) and F(2) are dual terms.

4. Dual term are separated by an index difference$\frac{I}{2}$ where,

$$I(l) \equiv 2^l \qquad [C.23]$$

5. If W_l^p is the weight at the Pth node, then the weight for the dual node is:

$$W_l^{(P+\frac{I}{2})} = - W_l^P \qquad [C.24]$$

6. The equation for computing any dual pair at the level l and position K:

$$X_l(K) = X_{l-1}(K) + X_{l-1}(K + \frac{I}{2})W^P \qquad [C.25a]$$

$$X_l(K + \frac{I}{2}) = X_{l-1}(K) - X_{l-1}(K + \frac{I}{2})W^P \qquad [C.25b]$$

The algorithm is so-called in-place algorithm with scrambled input and natural output. The terminology "in-place computation" is referred to that, after we simultaneously compute a dual node pair, we return the result to the storage location previously occupied by this dual node pair. The advantage of using the in-place computation is

that, the algorithm requires less total storage than natural input-output algorithm, but it requires that the input data must be scrambled before entering the computation array.

Starting from the program discussed, the author made several modification. One of them is to make 2DFFT program suitable for any size of rectangular input data. The other is to add filter subroutine for filtering purpose.

Since the size of image in our experiment is 256×128, M equals 256 and N equals 128.

There are mainly seven parts in the 2DFFT program. They are: MAIN program, subroutine DIM2FF, subroutine RFASF, subroutine CFASF, subroutine RSCRBLE, subroutine CSCRBLE and subroutine FILTER. The function of main program is to control all subroutines to make them work in proper order. It handles reading input data from input data file, writing output data into output data file and shifting the DC term in the spatial frequency domain to the array center. Subroutine DIM2FF can perform two-dimensional FFT either forward or inverse based on sign. When you are asked by this program, you can enter either +1 for forward 2DFFT or -1 for inverse 2DFFT. Subroutine DIM2FFT needs to call two subroutines. They are subroutine CFASF which performs column FFT. and subroutine RFASF which performs row FFT. Before performing one-dimensional FFT, we need to scramble input data. We use subroutine RSCRBLE and CSCRBLE to set up row and column scramble pointer vector. Finally, we use subroutine FILTER for filtering or modifying the intensity of image spectra. These routines are represented in following flowcharts. See Figure 12 on page 63 for MAIN program, Figure 13 on page 64 for subroutine DIM2FF, Figure 14 on page 65 for subroutine RFASF, Figure 15 on page 66 for subroutine RSCRBLE, and Figure 16 on page 67 for subroutine FILTER. The principle of subroutine CFASF is essential same with subroutine RFASF and subroutine CSCRBLE is essential same with subroutine RSCRBLE except the parameter.

Based on the principle of 2DFFT algorithm, we should able to expand this program to any size of rectangular input as long as we choose the sampling points in each direction as an integer power of two.

Finally, we would like to expand equation [C.6] and [C.7] from one-dimensional into two-dimensional.

$$\Delta X = \frac{L_x}{N} \qquad [C.26]$$

$$\Delta Y = \frac{L_y}{M} \qquad [C.27]$$

$$\Delta u = \frac{1}{L_x} = \frac{1}{N\Delta x} \qquad [C.28]$$

$$\Delta v = \frac{1}{L_y} = \frac{1}{M\Delta y} \qquad [C.29]$$

Where, the N is number of sampling points in X direction.

the M is number of sampling points in Y direction.

the L_x, L_y are the image dimensions in the X, Y directions.

the ΔX, ΔY are the sampling interval in the spatial domain.

the Δu, Δv are the sampling interval in the spatial frequency domain.

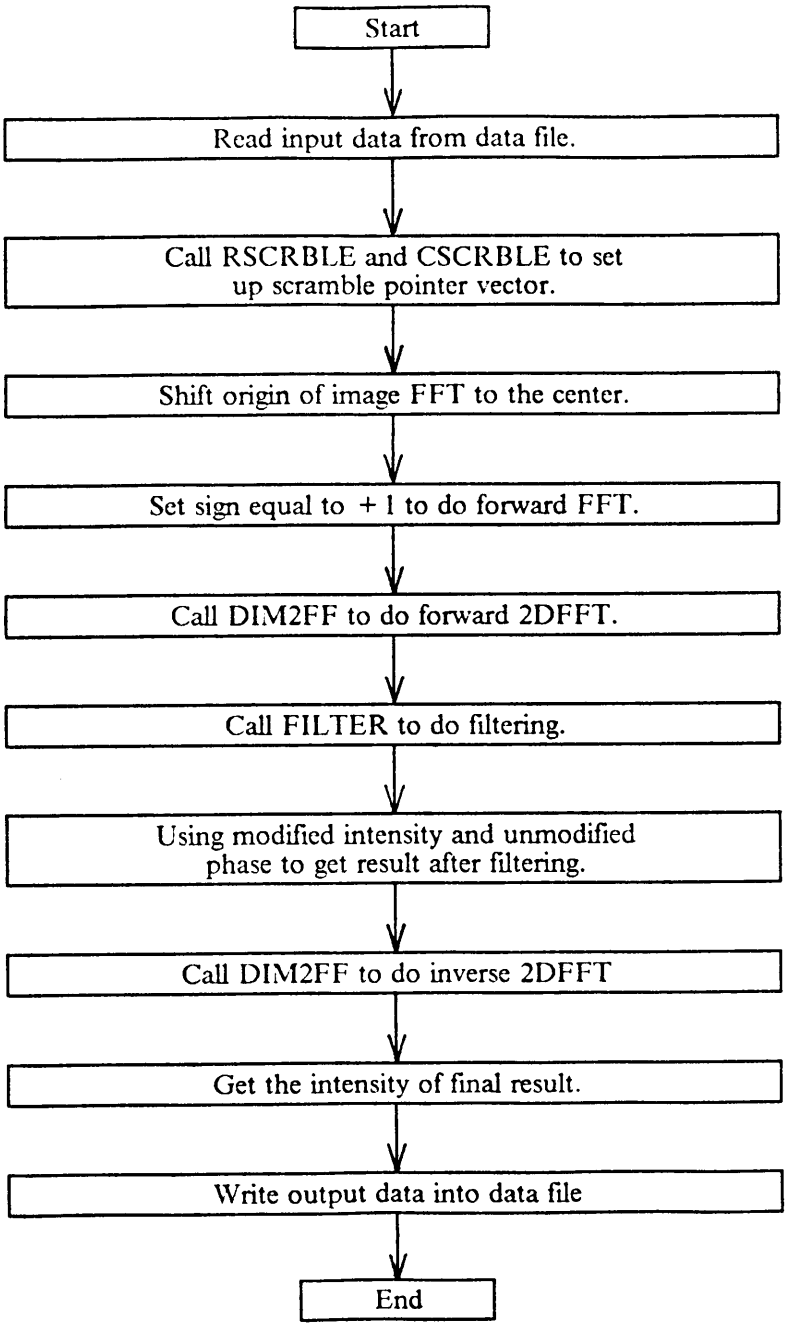


Figure 12. Flowchart for main program.

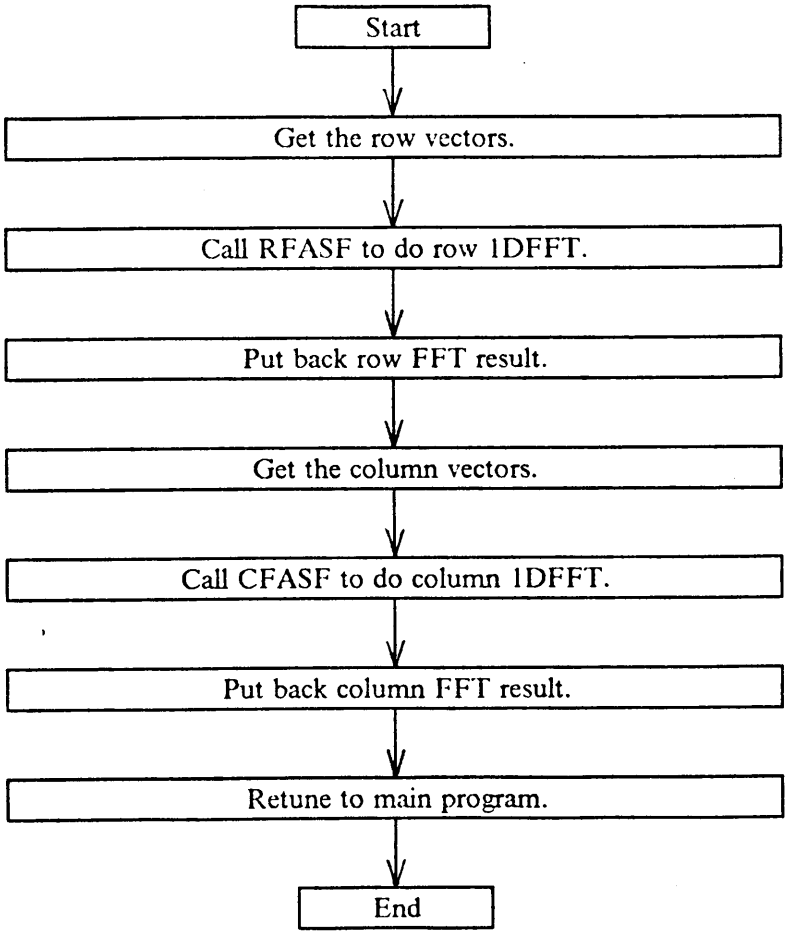


Figure 13. Flowchart for subroutine DIM2FF.

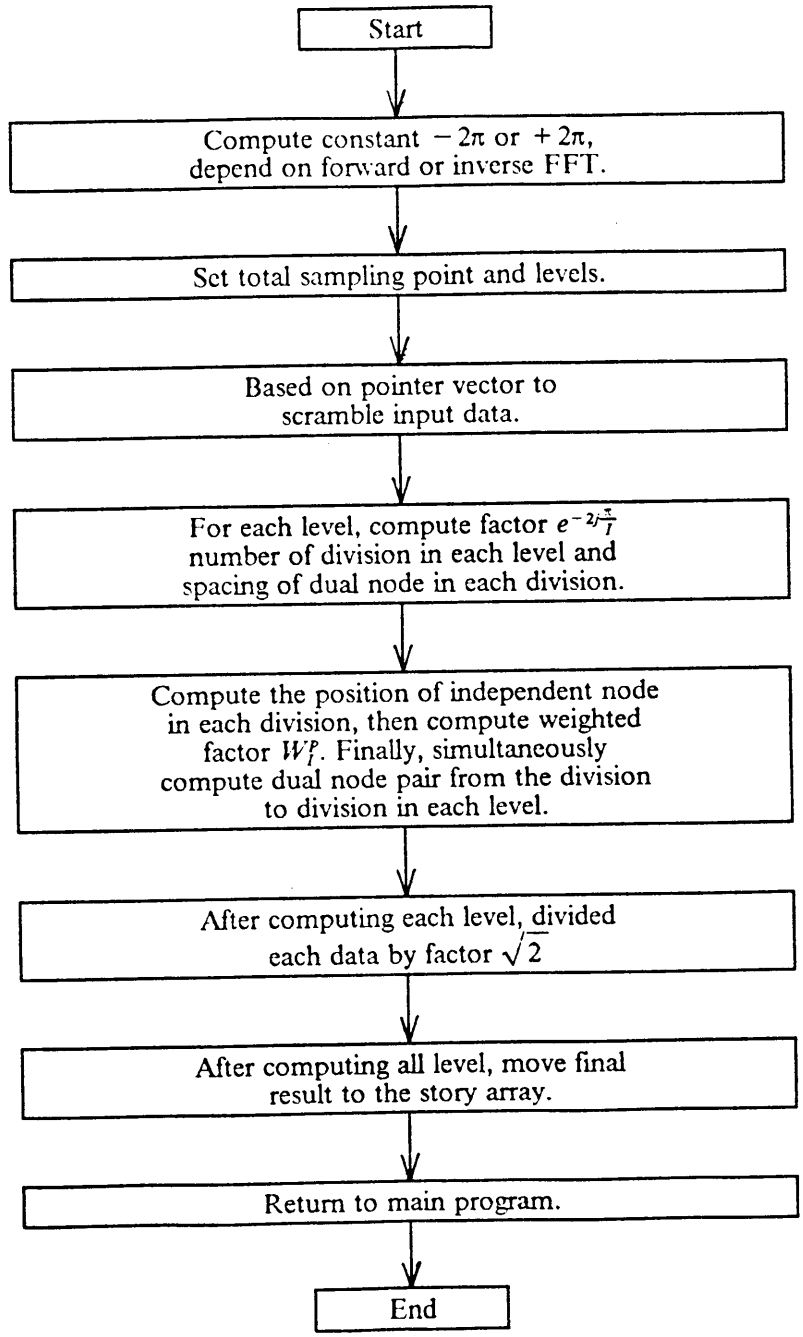


Figure 14. Flowchart for subroutine RFASF.

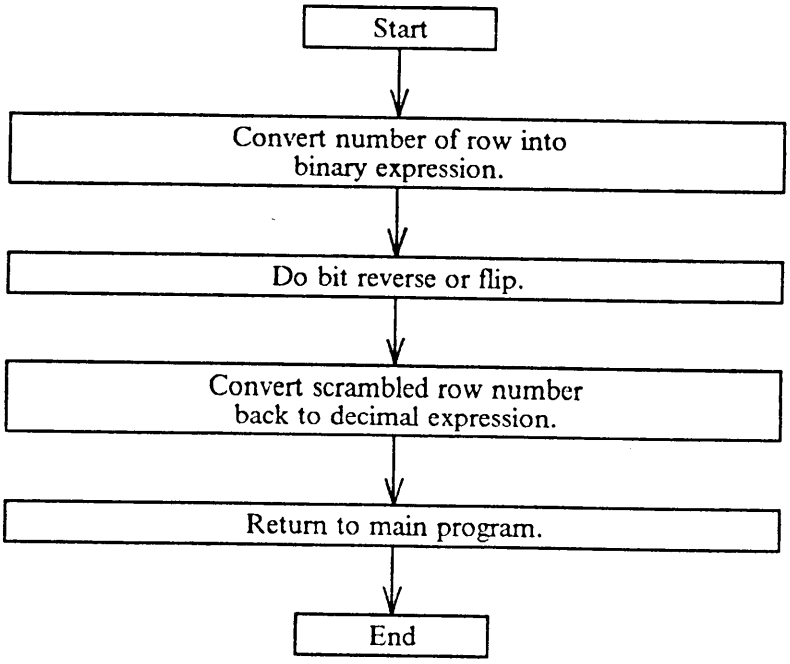


Figure 15. Flowchart for subroutine RSCRBLE.

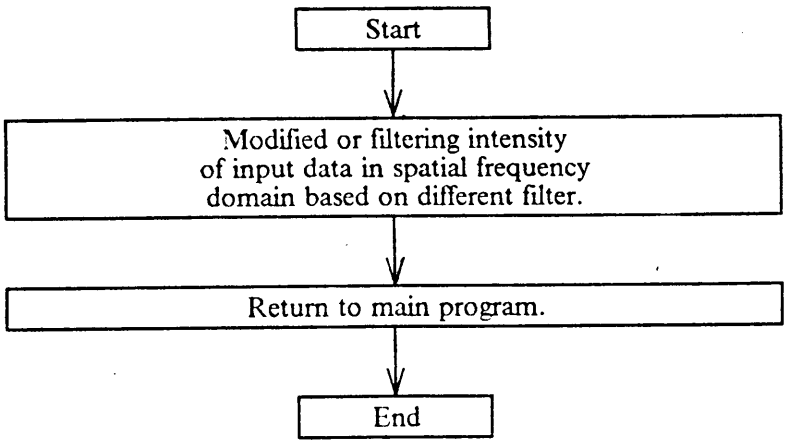


Figure 16. Flowchart for subroutine FILTER.

Appendix D. Extending the Depth of Focus Using Digital Image Filtering

G. H. Hu, R. J. Pieper, and T.- C. Poon

Department of Electrical Engineering
Virginia Polytechnic Institute and State University
Blacksburg, Virginia 24061

ABSTRACT

Several methods for extending the optical depth of focus using frequency domain operations are discussed. These methods are based on the intuitive concept that the spectral energy distribution of a focused image is biased toward lower frequencies after blurring. Examples are presented.

I. BACKGROUND

One of the most common factors contributing to the blurring is the physical separation between the detected image and the in-focus plane. This leads to a non-minimum circle of confusion [1].

Techniques for improvement in the optical depth of focus has been of current interest. An incomplete listing of these techniques follows.

The use of graded apertures has been shown to improve the optical transfer function and lead to a larger depth of field [2]. The use of an annular aperture combined with an optical equalization of the modulation transfer function have been shown to produce an improved depth of field [3]. A number of methods have been proposed for digitally analyzing a sequence of images each of which is composed of in-focus (sharp) as well as out-of-focus (blurred) regions [4,5,6]. The methods [4,5,6] are implemented without translating the spatial information to the frequency domain. The concept of employing frequency domain high pass filtering in order to produce edge sharpening [7] was introduced within the context of optical image processing. It was noted that removal of the frequency term of the spatial spectrum through the use of an opaque spot in the focal plane of the lens is analogous to inserting a capacitor in an electrical network. This operation optically has the effect of edge sharpening images.

We propose several digitally implemented methods based on a frequency domain interpretation of the effects of blurring. The examples provided are intended to demonstrate the feasibility of this approach.

II. THE IMAGE PROCESSING SYSTEM

The system as represented in Figure 17 on page 72 can be used to digitize, display and analyze images. With the aid of the computer, operations such as matrix addition, scalar multiplication and the more sophisticated fast Fourier transforms are readily implemented. Among the various FFT algorithms we employed the method using a bit reversed input producing a natural output [7]. The images to which the 2D-FFT were applied to had 256 $\times$ 128 pixels, (M x N). For reader convenience we include in the Appendix the essential discrete 2D Fourier transform relations that our FFT was based on.

III. DESCRIPTION OF THE METHODS EMPLOYED

In general we start with a number of images slices (see Figure 18 on page 73): each having distinct regions that are in focus and out of focus (blurred). On each slice a 2D FFT can be performed. The resulting frequency spectrum of each image slice can then be filtered. The details of the filter will in general depend on the specific spectrum characteristics of the image slice. For example, the spectrum characteristics can be represented by means of a histogram in the spatial frequency domain. Various forms of high pass filtering can be employed to deemphasize the low frequency blurred region relative to the crisper in-focus segments. The composite of the filtered transform can be inverse transformed according to Eq. (A1a). The remainder of this section is devoted to brief discussions of methods and results. Figure 20 on page 80 shows a set of blurred images to which our algorithms were applied. For comparison, Figure 20 on page 80 part d) is the in-focus image obtained by stopping down the camera lens (Figure 17 on page 72).

a) DC Suppression.

In this method the filtered spectrum $F'(u, v)$ satisfies

$$F'(u, v) = F(u, v) - F(0, 0) \tag{1}$$

where u and v are the spatial frequencies, and F(u,v) is the Fourier transform of the image slice. Operation (1) is applied to each of the blurred images. The three filtered spectrums are then added. It can be easily shown for discrete representations that removing the DC terms is equivalent to subtracting the average. For images having a dominant light background the average is close to the background level. Taking the inverse Fourier transform of the filtered composite spectrum results in the negative of Figure 20 on page 80 part e). In obtaining Figure 20 on page 80 part e). we performed a contrast reversal in the image as a last step. As can be seen from Figure 20 on page 80 part e). resultant image carries the objects spatial domain information over a greater depth range than any one of the blurred images and appears a bit more in focus than the space domain average of the blurred slices (Figure 20 on page 80 part h)).

b) Histogram Method.

For purposes of this method we define the average energy for the k^{th} frequency domain histogram band

$$E_k = \frac{1}{J_k} \sum_{m} \sum_{n} |F(u, v)|^2 , \qquad k = 1, 2, 3 \ldots \tag{2}$$
$$\text{in band K}$$

where J_k is the total number of pixels in the k^{th} band, and normalized average energy per band,

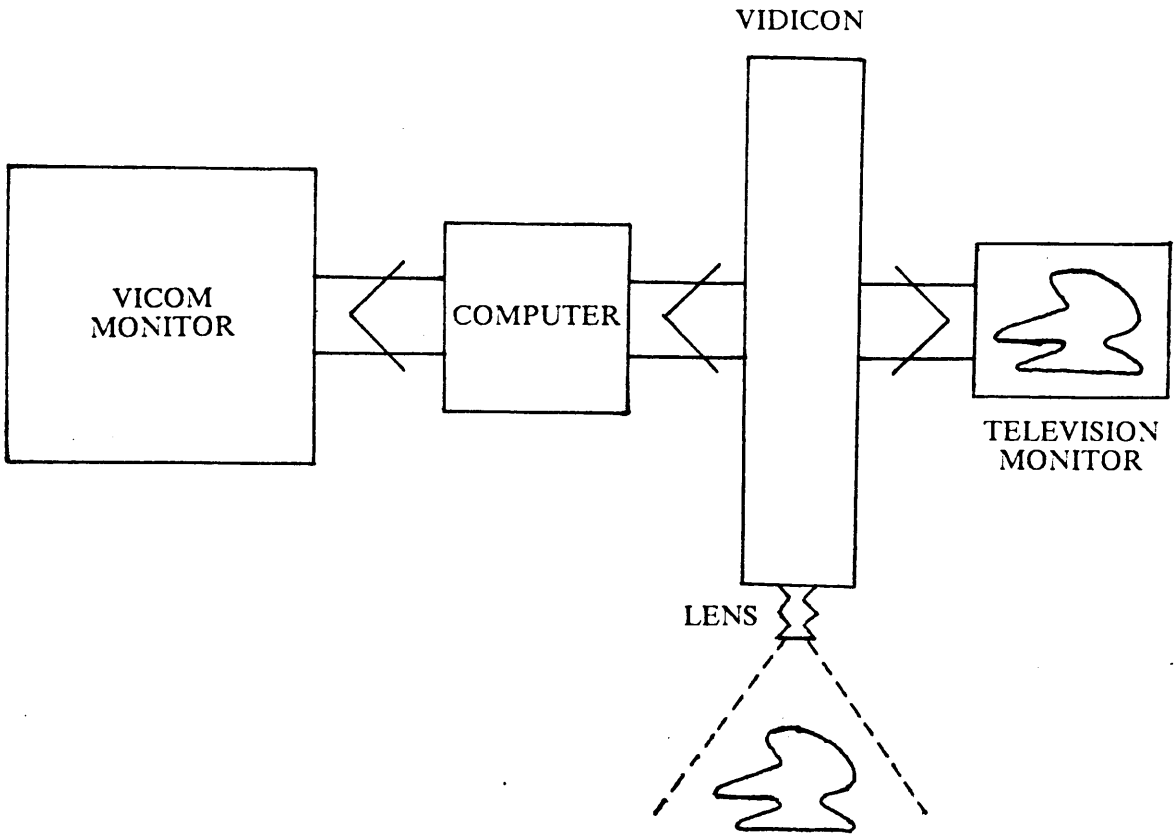


Figure 17. Digital image system.
The system used to digitize image slices and produce composite image.

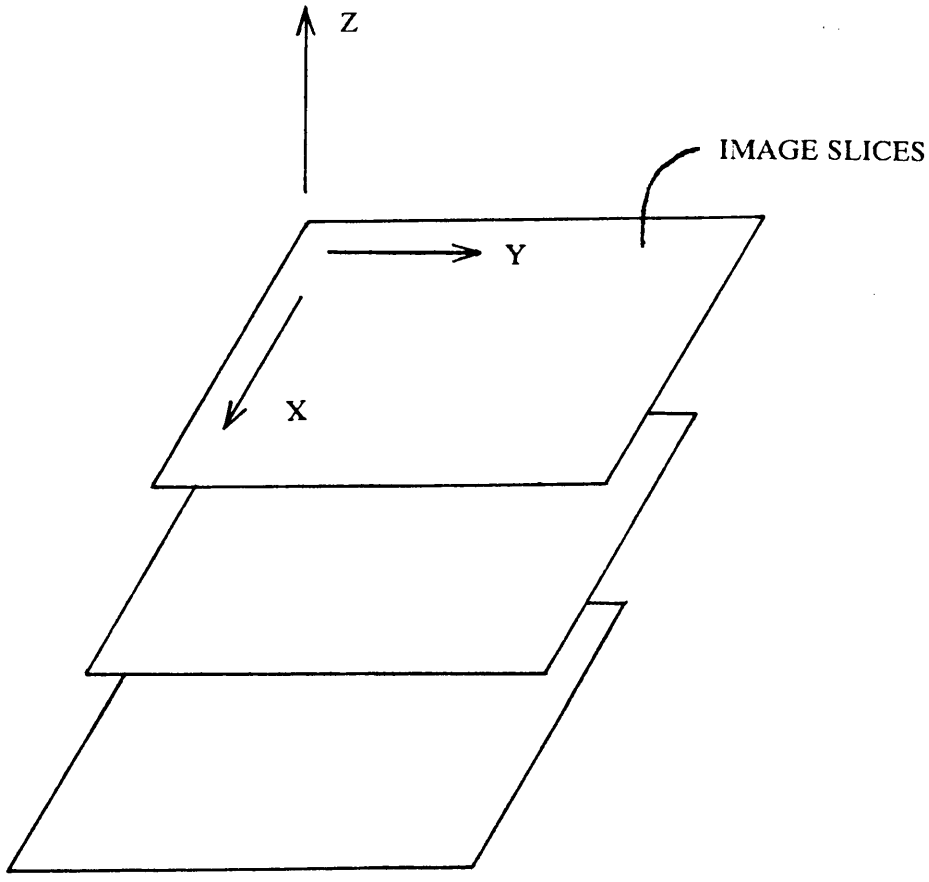


Figure 18. Sequential 2D image slices.

$$\varepsilon_k = E_k / \sum_{\substack{m \\ \text{all bands}}} \sum_{n} |F(u,v)|^2 , \qquad k = 1, 2, 3 \ldots \tag{3}$$

where the term in the denominator of Eq.(3) is, according to Parseval's theorem, equal to the total energy in the image. The post FFT filtering for each band is found from the ratio of the normalized average energy per band for in focus relative to the blurred images

$$H_k = (\varepsilon_k)_{Focus}/(\varepsilon_k)_{Blurred} , \qquad k = 1, 2, 3 \ldots \tag{4}$$

The filter H_k can be defined for any set of bands. We employed the following arbitrary band description requiring points within the k^{th} band to satisfy (See Appendix)

$$K^2_{(k-1)} < (u - \frac{N}{2} - 1)^2 + \frac{(v - \frac{M}{2} - 1)^2}{(\frac{M}{N})^2} \leq K^2_k \tag{5}$$

where specific values for the k indexed set K^2 are shown in Figure 19 on page 76. The result for this method is shown in Figure 20 on page 80 part f). One of the typical filters employed in this process has been sketched as shown in Figure 19 on page 76.

c) Low Frequency Suppression.

As can be observed from the filter in Figure 19 on page 76, the low frequency band k = 2 (the second band) tends to have significantly more energy in the blurred image as compared to the focus image, (i.e. $H_2 \sim .025$). This bias in energy can be contributed to blurring. In this method band k = 2 is completely suppressed. The results are shown

in Figure 20 on page 80 part g). For this case the results do not visually show significant improvement over the average (Figure 20 on page 80 part h).

Figure 20 on page 80 part j) shows the FFT amplitude spectrum of the in focus image and Figure 20 on page 80part i) is that of the blurred image (Figure 20 on page 80 part a). Close examination shows the frequency spectrum Figure 20 on page 80 part j) slightly more spread out.

IV. CONCLUSION

In two of the three methods tested we found modest improvement over the method of simply averaging. As a disclaimer the performance of these tests will depend to a significant degree on the characteristics of the object tested. We have demonstrated the feasibility of improving the image through various filtering methods.

APPENDIX

The following transforms relate the space domain representation f(x,y) to the frequency domain representation F(u,v):

$$f(x,y) = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{N}} \sum_{u=1}^{M} \sum_{v=1}^{N} F(u,v) e^{j2\pi ux/N} e^{j2\pi vy/M} \qquad (A1a)$$

$$F(u,v) = \frac{1}{\sqrt{M}} \frac{1}{\sqrt{N}} \sum_{x=1}^{M} \sum_{y=1}^{N} f(x,y) e^{-j2\pi ux/N} e^{-j2\pi vy/M} \qquad (A1b)$$

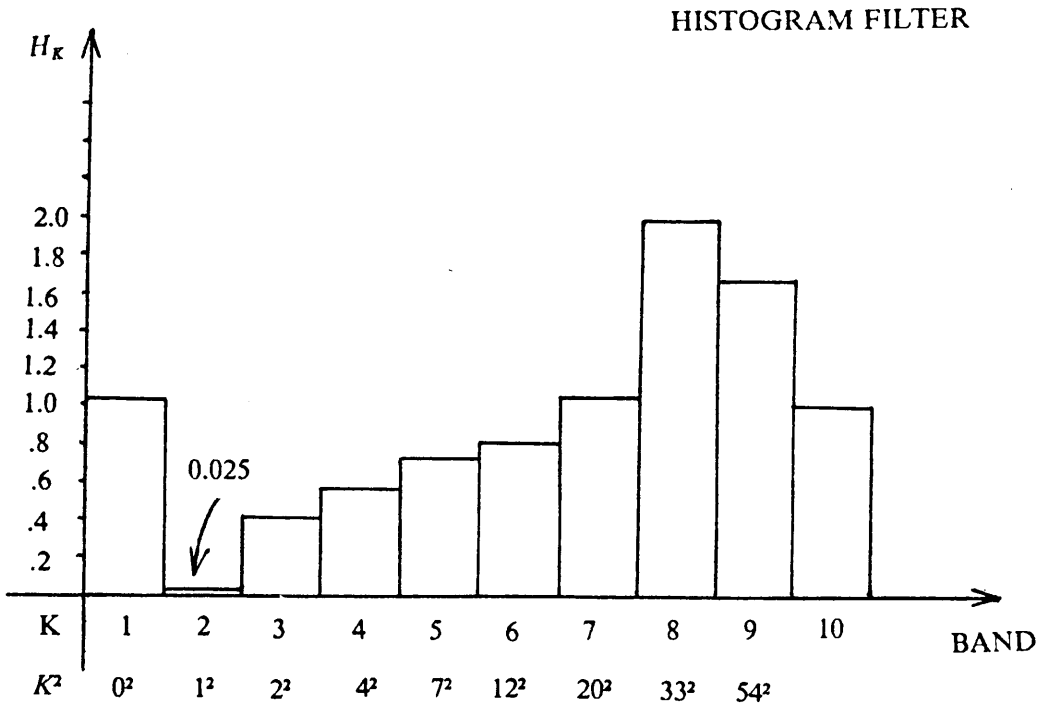


Figure 19. Filter used in histogram method.

where in our case N = 128, M = 256. This definition relates an ordered sampling of the image to an ordered sampling in the frequency domain of the sampled image. In order to interpret the indices x, y, u, v in terms of actual space and frequency coordinates the following can be defined:

$$\Delta X = \frac{L_x}{N} \tag{A2a}$$

$$\Delta Y = \frac{L_y}{M} \tag{A2b}$$

$$\Delta u = \frac{1}{L_x} \tag{A2c}$$

$$\Delta v = \frac{1}{L_y} \tag{A2d}$$

where ΔX, ΔY are the pixel dimensions and L_x, L_y are the image dimensions in the x,y directions respectively. In terms of these parameters the space coordinates of a pixel are given by,

$$(X\Delta x,\ Y\Delta y)\ x = 1,\ 2,\ \ldots,\ 128 \tag{A3}$$

$$y = 1,\ 2,\ \ldots,\ 256$$

The translation property of the Fourier transform can be applied to shift the DC point of the frequency domain (1,1) to the center pixel (65,129). The following are transform pairs [8]:

$$f(x,y)\ (-1)^{x+y} \leftrightarrow F(u - N/2,\ v - M/2) \tag{A4}$$

where N = 128 and M = 256. The shifted transform is obtained by applying (A1b) to the left side of (A4). After this transformation the maximum x frequency deviation represented will be equal to the maximum y frequency deviation since the pixel has square dimensions (i.e. $\Delta x = \Delta y$).

$$\frac{N}{2}\,\Delta u = \frac{N}{2}\,\frac{1}{Lx} = \frac{1}{2\Delta x} \qquad (A5a)$$

$$\frac{M}{2}\,\Delta v = \frac{M}{2}\,\frac{1}{Ly} = \frac{1}{2\Delta y} \qquad (A5b)$$

It follows that frequency domain surfaces of constant radial frequency satisfy:

$$[\Delta u\,(u - \frac{N}{2} - 1)]^2 + [\Delta v\,(v - \frac{M}{2} - 1)]^2 = (\frac{K}{L_x})^2 \qquad (A6)$$

which for $M \times N$ images form ellipses,

$$\frac{(u - \frac{N}{2} - 1)^2}{K^2} + \frac{(v - \frac{M}{2} - 1)^2}{K^2(\frac{M}{N})^2} = 1\ . \qquad (A7)$$

K^2 is an arbitrary positive constant which dictates the major and minor axis of the elliptical bands applied in the histogram method discussed in the text.

REFERENCES

1. E. Hecht and A. Zajac, Optics, Addison-Wesley Publishing Company, Dec. 1976.

2. M. Mino and Y. Okono, "Improvement in the OTF of a defocused optical system through the use of shaded apertures", Appl. Optics, vol. 10, no. 10, pp. 2219-2225, Oct. 1971.

3. J. T. McCrickerd, "Coherent processing and depth of focus of annular aperture imaginery", Appl. Opt., vol. 10, no. 10, pp. 2226-2230, Oct. 1971.

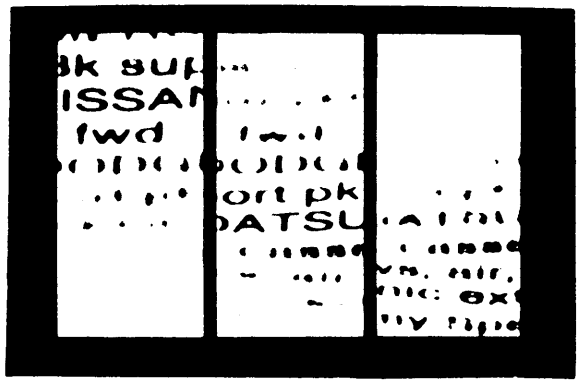
4. G. Hausler and E. Korner, "Expansion of depth of focus by 'image depuzzling'", Proceedings of 6th International Conference on Pattern Recognition, Oct. 1982.

5. R. Pieper and A. Korpel, "Image processing for extended depth of field", Applied Optics, vol. 22, no. 10, pp. 1449-1453, May 1983.

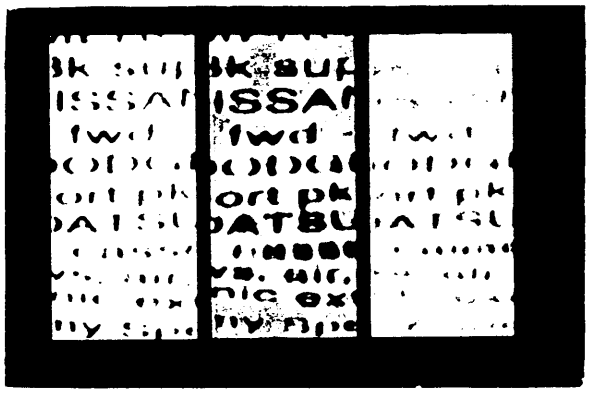
6. S. A. Sugimoto and Y. Ichiota, "Digital composition of images with increased depth of focus considering depth information", Appl. Optics, vol. 24, no. 14, pp. 2076-2080, July 1985.

7. E. L. O'Neill, "Spatial filtering in optics", IRE Trans. on Information Theory, pp. 56-65, June 1956.

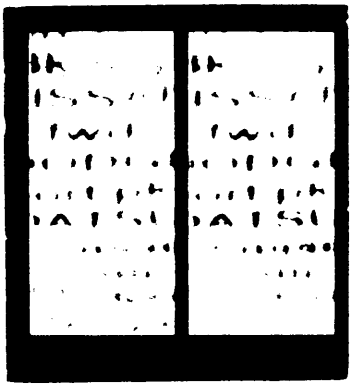
8. R. C. Gonzalez and P. Wintz, Digital Image Processing, Addison-Wesley Co., 1977.



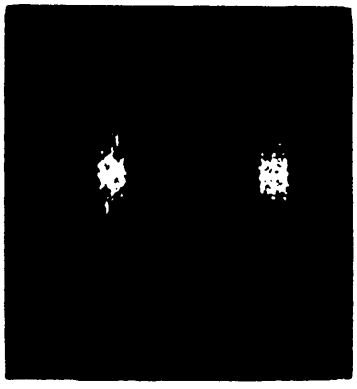
a. b. c.



d. e. f.



g. h.



i. j.

Figure 20. Input image set and result image.
a), b), c)Blurred image slices. d)In-focus. e)DC suppression. f)Histogram. g)Average. h)Low frequency suppression. Fourier amplitude spectrum. i)blurred image of a). j)In-focus image of d).

The vita has been removed from the scanned document