Spatial Enhancement

Nickolas Faust
The Electro-Optics, Environment, and Materials Laboratory
Georgia Tech Research Institute
Georgia Institute of Technology

Definition

- Spectral enhancement relies on changing the gray scale representation of pixels to give an image with more contrast for interpretation. It applies the same spectral transformation to all pixels with a given gray scale in an image. However, it does not take full advantage of human recognition capabilities even though it may allow better interpretation of an image by a user,
- Interpretation of an image includes the use of brightness information, and the identification of features in the image.
- Several examples will demonstrate the value of spatial characteristics in image interpretation.
- Spatial enhancement is the mathematical processing of image pixel data to emphasize spatial relationships. This process defines homogeneous regions based on linear edges.
- Spatial enhancement techniques use the concept of spatial frequency within an image. Spatial frequency is the manner in which gray-scale values change relative to their neighbors within an image. If there is a slowly varying change in gray scale in an image from one side of the image to the other, the image is said to have a low spatial frequency. If pixel values vary radically for adjacent pixels in an image, the image is said to have a high spatial frequency. Figure 1 (a-b) shows examples of high and low spatial frequencies:

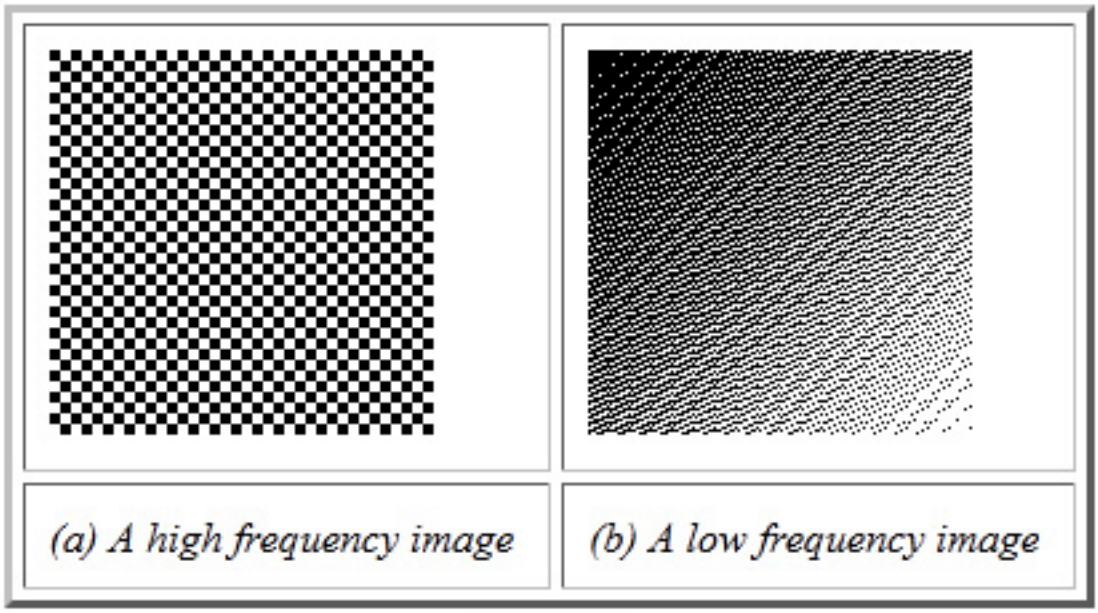


Figure1

- Many natural and manmade features in images have high spatial frequency:
  - Geologic faults
  - Edges of lakes
  - Roads
  - Airports
- Spatial enhancement involves the enhancement of either low or high frequency information within an image. Algorithms that enhance low frequency image information employ a "blurring" filter (commonly called a low pass filter) that emphasizes low frequency parts of an image while de-emphasizing the high frequency components. The enhancement of high frequency information within an image is often called edge enhancement. It emphasizes edges in the image while retaining overall image quality.
- Objectives or Purposes There are three main purposes that underlie spatial enhancement techniques:

To improve interpretability of image data

To aid in automated feature extraction

To remove and/or reduce sensor degradation Methods The two major methods commonly used in spatial enhancement are:

- Convolution
- Fourier Transform

- Convolution involves the passing of a moving window over an image and creating a new image where each pixel in the new image is a function of the original pixel values within the moving window and the coefficients of the moving window as specified by the user. The window, a convolution operator, may be considered as a matrix (or mask) of coefficients that are to be multiplied by image pixel values to derive a new pixel value for a resultant enhanced image. This matrix may be of any size in pixels and does not necessarily have to be square.
-
- Examples
- As an example of the convolution methodology, take a 3 by 3 matrix of coefficients and see the effects on an example image subset. A set of coefficients that is used for image smoothing and noise removal is given below:

1/9 *	1	1	1
	1	1	1
	1	1	1

3	3	4	4	5	6
2	3	3	4	4	5
1	2	2	3	3	4
1	1	2	4	4	7
1	2	4	20	20	20
2	3	6	20	20	20
2	3	4	20	20	20

If we have a sample image, given directly above where the image normally has a low smoothly varying gray scale, except for the bottom right region, which exhibits a sharp brightness change, we can see the effects of the convolution filter on a pixel-by-pixel basis.

Because we do not wish to consider edge effects, we will start the overlay of the

moving window on the x=2, y=2 pixel of the input image and end at the x=6, y=5 position of the original image.

The first p(x,y) (x=1, y=1), pixel of the output image would then be

Convolution Example

Examples
- As an example of the convolution methodology, take a
- 3 by 3 matrix of coefficients and see the effects on an example image subset.
- A set of coefficients that is used for image smoothing and noise removal
- is given below:

$$\frac{1}{9} * \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$$

If we have a sample image, given below:

3	3	4	4	5	6
2	3	3	4	4	5
1	2	2	3	3	4
1	1	2	4	4	7
1	2	4	20	20	20
2	3	6	20	20	20
2	3	4	20	20	20

where the image normally has a low smoothly varying gray scale, except for the bottom right region, which exhibits a sharp brightness change, we can see the effects of the convolution filter on a pixel-by-pixel basis.
Because we do not wish to consider edge effects, we will start the overlay of the moving window on the x=2, y=2 pixel of the input image and end at the x=6, y=5 position of the original image.
The first p(x,y) (x=1, y=1), pixel of the output image would then be

$$p(1,1) = \frac{1}{9} * \begin{matrix} (3*1 + & 3*1 + & 4*1 \\ + 1*1 + & 2*1 + & 2*1 \\ + 1*1 + & 1*1 + & 4*1) \end{matrix} \quad = 21/9 = 2.333$$

Because the output image, as well as the input image, is normally a whole number (integer) quantity, we will round the values to the nearest integer,

$$p(1,1) = 3.$$

Similarly,

$$p(1,2) = 1/9 * \begin{matrix}(3*1 + & 4*1 + & 4*1 \\ + 3*1 + & 3*1 + & 4*1 \\ + 2*1 + & 2*1 + & 3*1)\end{matrix} \quad = 28/9 = 3.111$$

$$p(1,2) = 3.$$

and

$$p(1,3) = 1/9 * \begin{matrix}(4*1 + & 4*1 + & 5*1 \\ + 3*1 + & 4*1 + & 4*1 \\ + 2*1 + & 3*1 + & 3*1)\end{matrix} \quad = 32/9 = 3.555$$

$$p(1,3) = 4.$$

Continued application of the same window (or filter kernel) will result in an output image given by:

2	3	3	4
1	2	3	4
1	4	6	9
2	6	11	15
3	9	14	20

This should be compared to the original data values for those pixel locations of

3	3	4	4
2	2	3	3
1	2	4	4
2	4	20	20
3	6	20	20

where there is a sharp discontinuity in the image.
The moving window filter, in effect, smoothed out the sharp discontinuity in the original pixel imagery.
A sample edge detection mask might be given as

-1	-1	-1
-1	8	-1
-1	-1	-1

and a value for p(1,1) would be

p(1,1) =	(-1*3)	+ (-1*3)	+ (-1*4)
	+ (-1*2)	+ (8*3)	+ (-1*3)
	+ (-1*1)	+ (-1*2)	+ (-1*2)

The resulting image after application of the mask is given by

4	-1	4	-2
1	-6	-2	-11
-7	-22	-26	-49
-4	-26	80	45
0	-28	46	0

assuming that only positive values are allowed in a image file, all values are offset by the absolute value of the minimum image elements (in this case +49).

The resultant image would then be:

53	48	53	47
50	43	47	38
42	27	23	0
45	23	129	98
49	21	95	49

Values greater than 90 are present in the output image and represent the edge of the bright region in the original image. Alternately, the negative values could be set to 0 giving and output image of

4	0	4	0
1	0	0	0
0	0	0	0
0	0	80	45
0	0	46	0

Again, the output images maybe compared to the original pixel values

3	3	4	4
2	2	3	3
1	2	4	4
2	4	20	20
3	6	20	20

One of the most used convolution kernels for edge enhancement of images was given by Chavez. The kernel is specified as:

$$1/9 * \begin{matrix} -1 & -1 & -1 \\ -1 & 17 & -1 \\ -1 & -1 & -1 \end{matrix}$$

Chavez originally derived the above kernel for enhancement of high frequency information in an ERTS MSS image. For a particular image pixel location and channel number, a low pass filter may be used to evaluate the average value in a 3 by 3 window. The convolution kernel would be given by:

$$avg = 1/9 * \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$$

The high frequency (HF) component in any given pixel wil then be given by

HF = pixel - *avg*

Represented in terms of a convolution kernel, this would be

$$HF = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{matrix} - 1/9 * \begin{matrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{matrix}$$

which means that

$$HF = 1/9 * \begin{matrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{matrix}$$

By adding the high frequency part, *HF*, back to the original pixel, a high frequency enhancement will be achieved:

New value = pixel value + *HF*.

This may be accomplished by:

New value

=

$$\begin{matrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{matrix} + 1/9 * \begin{matrix}-1 & -1 & -1\\ -1 & 8 & -1\\ -1 & -1 & -1\end{matrix}$$

Or

$$\text{New Value} = 1/9 * \begin{matrix}-1 & -1\\ -1 & 17\\ -1 & -1\end{matrix}$$

Figure 2 (a-b) shows two Landsat Thematic Mapper (TM) images of Downtown Savannah, Georgia. Figure 2(a) shows the area before enhancement, and Figure 2(b) shows the results of applying the above convolution kernel to the areas depicted in the image.

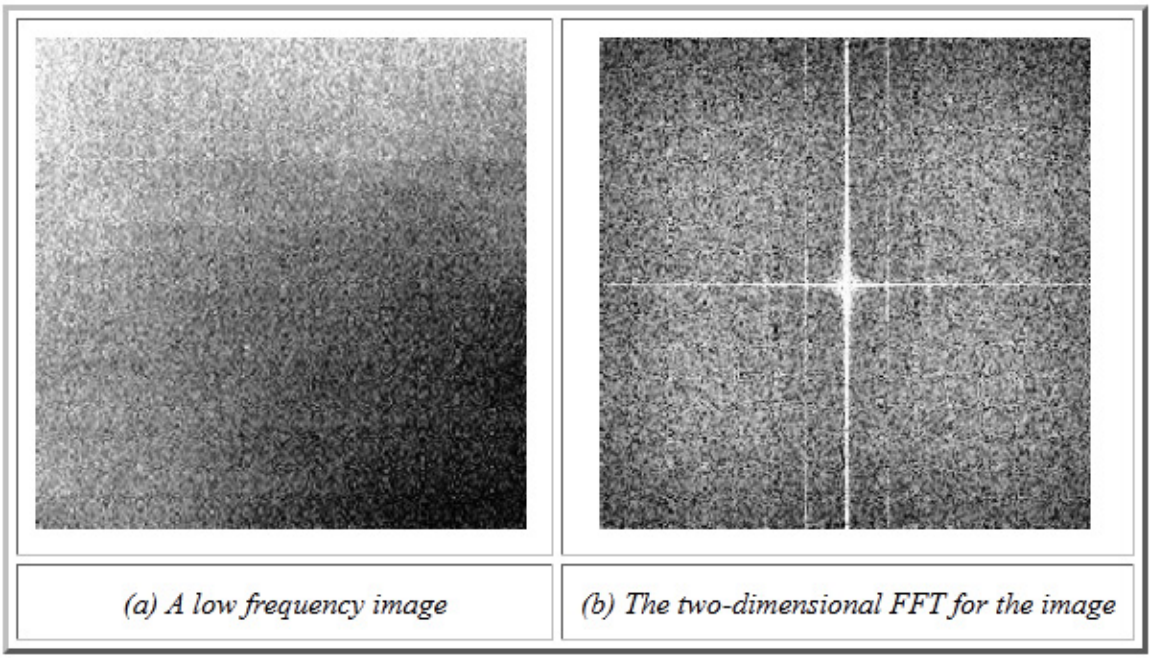
Fourier Transform Theory Definition

- Fourier transforms are used extensively in information theory, signal processing, and image processing.
- In the Fourier transform theory any one-dimensional function, f(x), may be fully represented by some superposition of trigonometric sine and cosine terms, F(x). The estimation of the coefficients and frequencies of each term necessary for full representation of the original function is involved in the calculation of the Fourier transform.
- Images are often discontinuous along a line or column, and end unceremoniously at the edges of the image. To handle this discrete image pixel data, a discrete version of the Fourier transform was developed and called the fast Fourier transform (FFT) due to Fourier transform theory assumes that the signal for

which the transform is desired is continuous with an infinite extent (ref. Cooley, Tukey).

- The performing of a two-dimensional Fourier transform on an image is equivalent to independently processing each single line of image data by a one-dimensional Fourier transform and then individually processing each single column of the results of the line-oriented one-dimensional Fourier transforms through another Fourier transform. This separability is a key factor in the implementation of two-dimensional transforms.
- Fourier transforms heavily utilize the theory of complex numbers and are often hard to visualize. By recalling examples at the first of the moving window convolution section, perhaps the interpretation of two-dimensional Fourier transforms will be made easier.
- Any image may be represented by a two-dimensional Fourier transform, which may be considered as an image with a real and a complex part. The two-dimensional FFT is a mapping of image pixel values into the image spatial frequency space. By performing a two-dimensional FFT on an image, we are creating a two-dimensional map of all spatial frequencies within an image.
- As a result of the FFT, every output image pixel has a real and an imaginary number associated with it. The real pixels form an image that may be thought of as the magnitude of the spatial frequencies present in an image, and the imaginary pixels form and image representing the phase of the spatial frequencies. As shown above, the highest spatial frequency that can be present in an image is equivalent to every other pixel having black-and-white values. Therefore, if an x and y axis are used to represent spatial frequencies on a plot, the width of the plot will, at most, be the total width of the image divided by 2.
- A useful way to display the spatial frequencies within an image is by using a star diagram representation of the magnitude of the complex two-dimensional FFT. In such a diagram, the lowest frequency component within an image (the average value, or albedo, of the image) is shown at the center of the diagram, and spatial frequencies increase pixel by pixel away from the center of the diagram. The brightness of the pixels at each x and y position relate to the relative occurrence of that spatial frequency in the original image.
- Spatial frequencies only exist up to the Nyquist frequency in the x and y directions, so the display is reflected about the center of the diagram. Furthermore, information in the +x and +y direction from the diagram center duplicates information in the -x and -y direction of the diagram. Figure 3(a) is

the same image shown in Figure 1(b). Figure 3(b) shows the magnitude of the two-dimensional FFT for the same image. Note that the majority of the spatial information(bright values) in the two-dimensional FFT is in the lower frequencies, as indicated in the original image.



(a) A low frequency image

(b) The two-dimensional FFT for the image

Figure 3

Figure 4(a) depicts the same checkerboard image shown in Figure 1(b), and Figure 4(b) shows the two-dimensional FFT for the image. The magnitude image has high values along 2 lines crossing in the center of the star diagram. The outside edge of the star diagram also has high values showing an abundance of high frequency information.

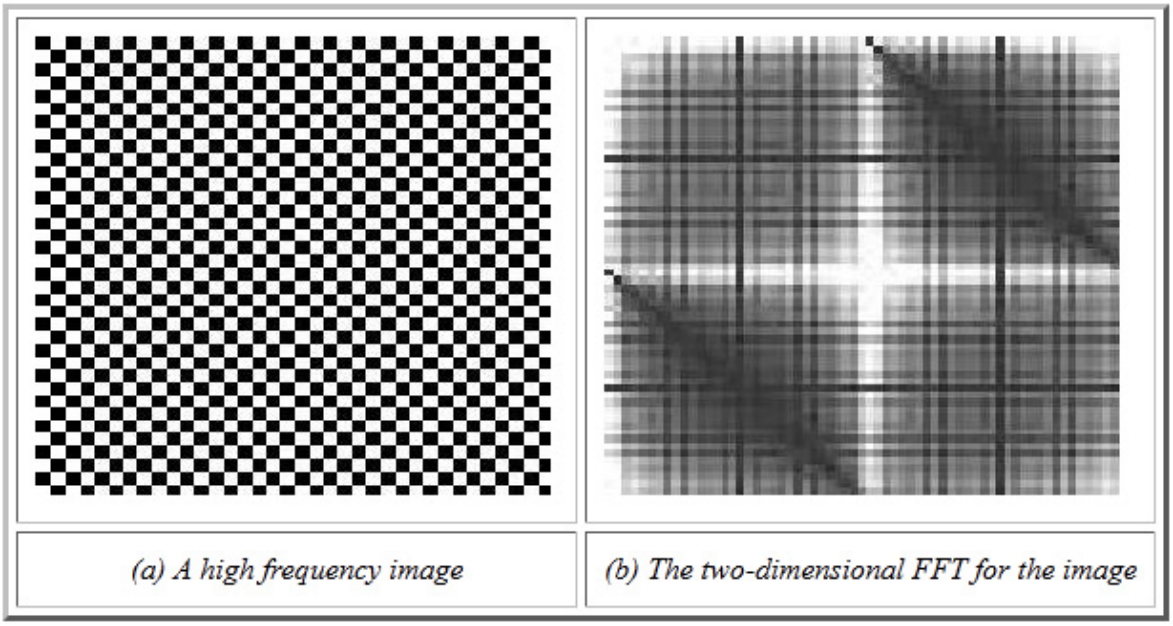


Figure 4

Fourier Transforms and Image Enhancement

Fourier Transforms and Image Enhancement

A two-dimensional FFT image may be useful in itself in developing an understanding of individual images, but Fourier Transform theory lends itself to image enhancement techniques as well. The ability to produce a two-dimensional FFT star diagram is known as the running of a *forward*

FFT. This process can also be thought of as transforming an image from the normal *time domain* to the *frequency domain*. The resulting frequency domain image may be transformed back to the time domain by performing an *inverse two-dimensional FFT*.

If no changes are made to the spatial frequency complex image, the inverse two-dimensional FFT will provide the exact same image that we began with. Fourier theory, however, tells us that we may perform certain operations, called *convolutions,* in the frequency domain that may enhance the image after the inverse two-dimensional FFT. These frequency convolutions are not to be confused with the

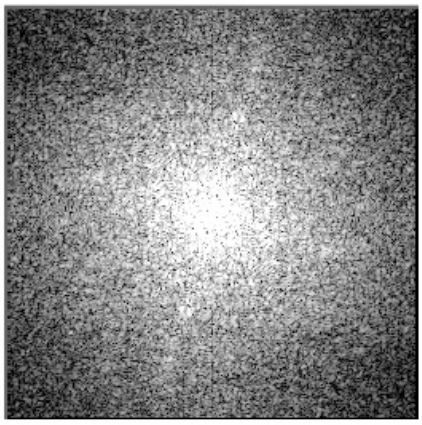
kernel convolutions discussed above.

A convolution in the frequency domain is a simple multiplication of an image mask that may be arbitrarily designed by a user, multiplied by the complex frequency domain image. The resultant frequency domain image is then run through the inverse two-dimensional FFT process to yield a transformed image.

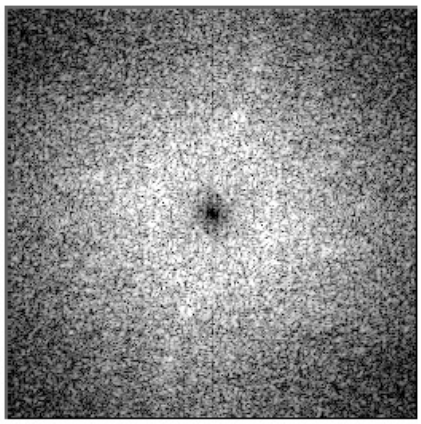
This process of convolution in the frequency domain is extremely valuable in the spatial enhancement of image data. We may perform the operations discussed earlier with kernel convolution in a more complete and flexible manner. In addition, there are some functions that may be done by frequency convolution that as yet have not been achieved by kernel convolution, such as *noise removal* from an image, and *image restoration*

An example of high-pass filtering is shown below, using the Savannah dataset, in Figure 5(a-d). To create a high-frequency enhanced image, the high spatial frequency components of the image are extracted and added back to the original image. This is easily done using frequency convolution. First, a TM image (Figure 5(a)) is transformed into the frequency domain (Figure 5(b)). Next, a mask is developed in the frequency domain, which is 0 for all spatial frequencies less than the selected value and 1 for all spatial frequencies greater than the value (Figure 5(c)). Thus, only the high frequency parts of the complex spatial frequency image are retained. When the inverse two-dimensional FFT is performed, the resultant image represents a high pass filter of the original image (Figure 5(d)). It is simple to define masks to be used in the frequency domain, but one must be careful to know what types of effects to expect in the time domain.

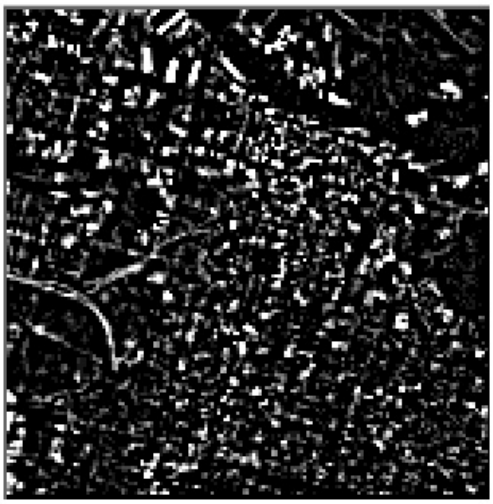
(a) Band 1 Landsat TM image of Downtown Savannah



(b) The two-dimensional FFT for the image



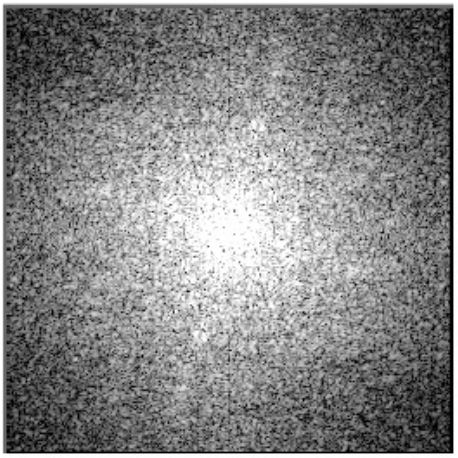
(c) The high pass mask used on the FFT



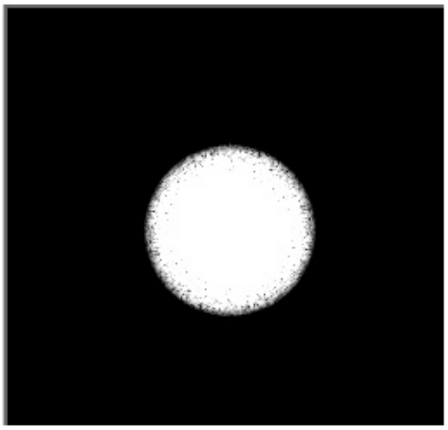
(d) The resultant image

An example of low-pass filtering using Fourier Transforms is shown below, in Figure 6(a-d):

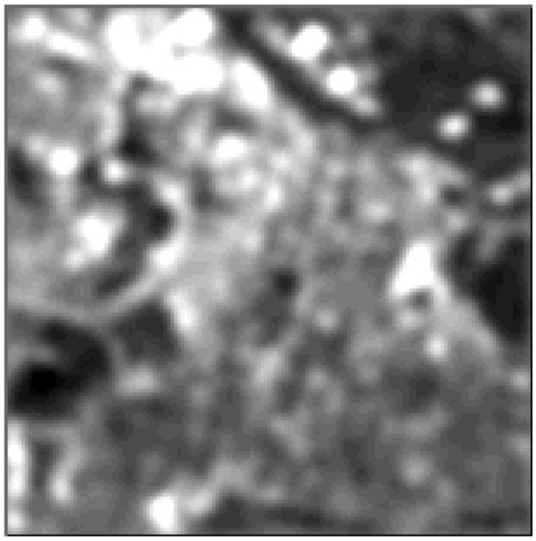
(a) Band 1 Landsat TM image of Downtown Savannah



(b) The two-dimensional FFT for the image



(c) The low pass mask used on the FFT



(d) The resultant image

Notes P.S. Chavez, Jr., "Atmospheric, Solar, and MTF Corrections for ERTS Digital Imagery," in Proceedings of the American Society of Photogrammetry Fall Meeting, October 1975, p. 68.