DIGITAL FILTERING TECHNIQUES FOR X-RAY IMAGE ENHANCEMENT

E. L. HALL
Departments of Electrical Engineering and Radiology
University of Missouri-Columbia

Summary Image enhancement techniques are designed to improve image quality for human viewing and are especially important in X-ray imaging systems since radiation levels must be minimized. Recursive spatial digital filtering and gray level transform enhancement techniques which may be implemented with small memory requirements and fast computation times will be described. A comparison of computations shows that this method is competitive with fast transform methods in terms of the number of computations and advantageous in terms of memory requirements. The most common method for evaluating enhancement results is subjective human evaluation. An experimental comparison of subjective evaluations and some objective measurements for a series of chest X-rays illustrate the evaluation problem and a possible solution.

Introduction An X-ray image is produced by permitting an X-ray source to penetrate an object and expose a photographic film. The intensity of exposure at a particular point on the film depends primarily on the object density and the source-object-film geometry. The X-ray tube has a finite focal spot which limits resolution. Also, phosphor intensifying screens are employed to hasten film exposure.

The information content of a radiograph depends upon the absorption characteristics of the object, which are determined by the thickness, density, and material properties of the object exposed at a selected X-ray energy level. The resolution of a radiographic recording system depends primarily upon the X-ray focal spot size [1], the speed of the image intensifier screen(s) used [2], object scatter and motion, and the object-film distance. X-ray image contrast is affected primarily by kilovoltage rating, and in a secondary way by object scatter. In addition, the random effect of quantum mottle, resulting from spatial changes in X-ray illumination, adds independent noise to the image. Provided that the combination of milliamperage and the source-object distance is adjusted so that sufficient film exposure is available and object-film distance is minimized, the image contrast and resolution depends primarily on the kilovoltage and focal spot size, respectively.

In general radiographs are characterized by their low resolution and low contrast ratios of small features superimposed onto uniform backgrounds. Digital techniques for image enhancement have been used by several researchers [3-5]; however, further developments in on-line digital enhancement and methods for evaluation enhanced images are necessary in many applications.

The purpose of this paper is to describe recursive spatial digital filtering and gray level transformation techniques for image enhancement and illustrate an approach to the problem of objectively evaluating image enhancement results.

First, a distribution linearization technique for contrast enhancement will be described. Next, spatial digital filtering by recursive partial difference equations will be considered. Finally, a theoretical and experimental evaluation of the quality of several processed chest X-rays will be described.

Distribution Linearization Because image digitalization requires that each gray level value be quantized over a finite range (e.g., 6 bits/ picture element), a histogram of the gray level distribution values may be easily computed.

The gray level histogram (first order probability density function) of most images exhibits a brightness peak which is heavily biased to the dark side of the histogram. The digital representation of the radiographs in this paper have 256 X 256 squared brightness values over a 33 X 33 squared centimeter face, with a sampling frequency in either dimension of .78 lines/mm. The effect of having a large percentage of the 65536 picture elements (pixels) concentrated into such a narrow portion of the histogram is an image with little visible detail. Many contrast ratios that define edges are simply not displayable. One technique frequently used to correct this is the position invariant, nonlinear application of a software or hardware logarithmic conversion of the brightness pixels. Because a more rectangular histogram is advantageous, a position invariant, nonlinear histogram equalization technique, as described in (6) is useful. Figure 2 is the result of linearizing the distribution of the image shown in Figure 1. Also, Figure 7 is the result of linearizing the distribution of the image shown in Figure 5.

Recursive Spatial Digital Filtering Spatial frequency domain filtering, performed either digitally, optically, or electronically, is a useful technique in several image processing applications. The usefulness is exemplified by the many applications which include image enhancement, restoration, template matching, and noise filtering. Digital filtering techniques avoid optical bench and electronic drift requirements but require image digitization, storage and computation availability.

Figure 8 is the magnitude of the Fourier transform space of the image shown in Figure 1. The center value (zero frequency coefficient) contains the image d.c. level (average

image gray level). The higher energy, lower frequency components, radially surrounding the center, contain much of the image contrast information. The low energy, higher frequency coefficients are primarily responsible for image edge information. The transform of Figure 8 is indicative of most two-dimensional Fourier transforms. The sampling frequency in this example is .78 ℓ/mm over a square image. Thus the maximum frequency on both axes is .39 ℓ/mm.

The dichotemy of the spatial frequency domain information is clearly illustrated by low pass and high pass spatial filtering as shown in Figures 2 and 3. The low pass filtered image contains the overall contrast information while the high pass filtered image contains the edge and detail information. A filter which is useful for image enhancement is a high frequency emphasis filter.

The cross section of a Butterworth high frequency emphasis filter is shown in Figure 9. Note that the filter retains some contrast information while emphasizing the higher spatial frequencies. At normal viewing distances, the high emphasis frequency response matches the response of the human visual system. A high emphasis filtered image computed via the fast Fourier transform (FFT) with an isotropic filter is shown in Figure 5. The versatility of the FFT and economy of computations makes the method ideally suited for spatial filtering design experiments; however, for a large class of image enhancement situations, a large computer or array processor is not available or economically feasible; for example, on-line medical X-ray imaging for clinical diagnosis or industrial X-ray imaging for quality control. As will be shown, the recursive spatial digital filter is suited to these applications.

The three basic computational algorithms for implementing spatial frequency domain digital filters are based on direct convolution, the fast Fourier transform, and recursive partial difference equations.

The general form of the digital imaging computation is shown in Figure 10, and given by:

$$g(nX,mY) = \sum_{k=0}^{K_1-1} \sum_{\ell=0}^{K_2-1} a_{k\ell} f(nX-kX, mY-\ell Y) + \sum_{k=0}^{L_1-1} \sum_{\ell=0}^{L_2-1} b_{k\ell} g(nX-kX, mY-\ell Y)$$

$$\{k + \ell \neq 0\}$$

Assuming that the system is causal, that is, f(nX,mY) = 0 and g(nX,my) = 0 for n,m < 0, one may use the shifting theorems of the ZW-transform to develop the transfer function for the system. Therefore, assuming $b_{OO} = 1$:

$$H(z,w) = \frac{G(z,w)}{F(z,w)} = \frac{\sum_{k=0}^{K_1-1} \sum_{\ell=0}^{K_2-1} a_{k\ell} z^{-k_w-\ell}}{\sum_{k=0}^{L_1-1} \sum_{\ell=0}^{L_2-1} b_{k\ell} z^{-k_w-\ell}}$$

A method used to describe the recursive filter operation is a flow diagram, as shown in Figure 11. The input and output image arrays are indicated by small circles. Lines connect certain image points to a summing junction. Coefficient multiplication is indicated by the lower case letter next to a line. Note that the diagram clearly indicates that each output array value is a linear combination of previous and present inputs and previous outputs.

Most image input devices are sequential rather than random access. For example, a scanning system or digital tape produces or stores data in a point-by-point, line-by-line manner. The only internal memory storage requirement is space for the filter input and output buffers which would consist of ($L_1 L_2 + K_1 K_2$) locations. A method for shifting the data for each row point and each column point is also necessary. An arithmetic unit which computes the additions and multiplications and accumulates the sum is also required. Finally, the output may be stored on another sequential device, such as a tape or disk file. Since the filter computation is simple and requires only a small number of memory locations, it is ideal for small computer image processing applications.

The three computational algorithms will now be compared. Although each method is defined by a convolution relation, the actual convolution computation for each method is different. The normal convolution computation produces an aperiodic, equal point convolution. The FFT method produces a periodic full range convolution. The recursive equation produces an aperiodic, variable range convolution. Thus, a comparison of the three methods is somewhat complicated.

The core memory requirements for the three methods will also be compared.
The first computational method, aperiodic convolution, is described by:

$$g_1(nX,mY) = \sum_{i=0}^{M_1-1} \sum_{j=0}^{M_2-1} h_1(iX,jY) f)nX-iX,mY-jY)$$

where g1(nX,mY) is the filter output and f(nX,mY) is the filter input, which is defined for:

$$n = 0, 1, \ldots, N_1-1\ldots$$
$$m = 0, 1, \ldots, N_2-1\ldots$$

and $h_1(iX,jY)$ is the filter point spread function, which is non-zero for:

$$i = 0, 1,\ldots M_1-1$$
$$j = 0, 1,\ldots M_2-1$$

and zero, otherwise. Usually, the filter point spread function is much smaller than the image frame, i.e.:

$$M_1 < < N_1$$
$$M_2 < < N_2$$

The number of computations required for filtering an N_1 x N_2 point image is approximately:

$$NC_1 = M_1M_2\ N_1N_2$$

real operations where a real operation is defined as a real multiplication and addition.

The periodic convolution which is equivalent to FFT transform multiplication and inverse transformation is given by:

$$g_2(nX,mY) = \sum_{i=0}^{N_1-1} \sum_{i=0}^{N_2-1} h_2(iX,jY)f\{[n-i]X,[m-j]Y\}$$

where

$$[n-i] = (n-i) \bmod N_1$$
$$[m-j] = (M-j) \bmod N_2$$

The modulus expressions indicate the periodic or “wrap around” nature of FFT equivalent convolution. Also note that the computation is over the entire frame. The number of computations using an FFT algorithm for computing a transform is proportional to $2N_1\ N_2\ \log_2 N_1N_2$ complex operations where a complex operation is defined as a complex multiplication and addition. Therefore, if the filter operation consists of computing the transforms of the input function, point-by-point multiplication of the input by a filter transform, and inverse transforming the result, the number of computations required is approximately:

$$NC_2 = 4\ N_1N_2\ \log_2 N_1N_2 + N_1N_2$$

The final computational algorithm produces an equivalent convolution of the form:

$$g_3(nX,mY) = \sum_{i=0}^{n} \sum_{j=0}^{m} h_3(iX,jY) f(nX,mY-jY)$$

If the output image frame was computed by the above equation, the number of computation required would be:

$$\sum_{i=0}^{N_1-1} \sum_{j=0}^{N_2-1} (i+1)(j+1) = \left(\frac{N_1^2+N_1+2}{2}\right)\left(\frac{N_2^2+N_2+2}{2}\right)$$

However, the recursive filter computation requires only

$$NC_3 = (K_1K_2 + L_1L_2) N_1N_2$$

real operations.

In order to compare the three algorithms one must assume equal impulse responses. It is useful to consider two cases corresponding to small and large non-zero convolution areas. For simplicity, one may assume square areas, i.e.:

$$K_1 + K_2,\ L_1 = L_1 = L_2,\ M_1 = M_2,\ N_1 = N_2$$

Case 1. Small Convolution Area

Suppose the filter responses are equal, i.e.:

$$h_1 = h_2 = h_3 = h$$

and the convolution area is small, i.e.

$h(nX,mY) = 0$ for $n,m > M$

The number of computations by direct convolution is:

$NC_1 = M^2N^2$ real operations

while the number of computations for FFT filtering is:

$NC_2 = N^2 \{8\log_2 N+1\}$ complex operations

Finally, the number of computations for the recursive equation is:

$NC_3 = (K^2+ L^2)N^2$

However, since the impulse response has been constrained to be finite, the filter has essentially been constrained to be non-recursive, thus:

$$K = M$$

$$L = 0$$

Thus, for this case, the direct convolution and non-recursive filter are identical and the computation trade-off is simply the well-known comparison between direct convolution and the FFT. [7]

Case 2. Large Convolution Area

Again assume the filter responses are equal, but suppose the nonzero response is large, i.e.:

$$h(nX,mY) \neq 0 \text{ for } n,m \leq N$$

For this case, the number of computation required by direct convolution is:

$$NC_1 = 4$$

while the FFT requires:

$$NC_2 = N^2 \{8 \log_2 N+1\}$$

However, the number of computations required for the recursive filter is still:

$$NC_3 = (K^2 + L^2)N^2$$

For large values of N, clearly NC_2 and NC_3 will be much smaller than NC_1. Thus, the main comparison is between the M and recursive filter. Using the conservative estimate that one complex operation is equivalent to two real operations permits one to develop the situation in which recursive filters are advantageous as:

$$\frac{K^2 + L^2 - 2}{16} < \log_2 N$$

If K = 1 and N = 256, the above equation requires $K^2 < 65$.

Thus, a recursive filter with 32 feed forward and 32 feed back coefficients would still be advantageous. Therefore, for the case in which the convolution area is large, both the FFT and the recursive filter are much faster than direct convolution. Furthermore, since

K^2 and L^2 may not increase as fast as $\log_2 N$, the recursive filter is at least competitive and may have a computation advantage over the FFT.

It is useful to note that both the FFT and the recursive filter compute a different convolution than the aperiodic convolution. However, the "growing" type response is usually not as bothersome in image processing as a periodic "wrap-around".

The amount of core memory required to perform a computation is also an important consideration, especially in real time digital image processing. Both aperiodic convolution and recursive filters require only a small core memory when they are computationally advantageous. That is, for the small filter response case, M^2 is small. For the large filter response case, $K^2 + L^2$ is also small. Thus, both methods are suitable for small computers and real time computations. For a single pass FFT computation, a core memory large enough to accommodate the image array, i.e. N^2, is required. Large array transforms can only be computed in real time with a large computer or a special purpose FFT processor.

Evaluation of Enhanced Images The most effective method of comparing images at present is subjective human evaluation. However, the inconsistencies encountered motivate the search for a more objective measure.

An experimental comparison of subjective evaluations of radiologists and bioengineers and several objective measures of quality will now be described. The comparison illustrates the differences in subjective evaluations by groups of observers and illustrates the amount of agreement between objective and subjective evaluations.

The method of evaluating image quality is dependent on the processing application. For an image communications application, the ideal image is assumed transmitted and a degraded image is received. Rate distortion theory provides a theoretical solution to this evaluating problem in that a weighted square error distortion measure may be selected and a bound on the number of bits/picture element necessary to satisfy the distortion level may be determined.

Spectacular examples of image enhancement and restoration are frequently encountered. For example, features which are barely visible in an original image may be easily visible in an enhanced image. Similarly a blurred image may be sharpened by spatial restoration filtering. However, the more general enhancement and restoration problems present a difficult evaluation problem. The difficulty is due to the lack of a standard reference image. For example, an original image is presented and some improvement is desired but the ideal improvement is not known.

Several factors influence subjective image evaluation. Contrast information is important not only in the number and linearity of distinguishable gray levels but also in the overall contrast ratio. The distinguishability of objects depends on the contrast levels in the objects, the frequency content of the object, and the frequency response of the human visual system.[8,9] Resolution is a measure of the number of distinguishable spatial regions in an image frame, and provides an upper bound on system performance. However, acuity or sharpness in an image is related to rendition of objects much larger than those barely visible. Sharp edge rendition on large objects may be an expected or learned performance. Noise effects are also important. Finally, it is possible that the HVS response is not isotropic.

These factors affect human subjective evaluation and may be studied individually. However, the overall effect of these factors on subjective evaluation is a complex psycho-physical relationship that is not understood. For the experimental comparison, evaluator groups of five radiologists and five bioengineers were asked to rate a series of processed images on a 0 to 100 scale. The average rating for each group is shown in Figure 12. Note that the abscissa numbering corresponds to the image number in Figures 1-7. The two groups agree on most of the images but differ noticeably on the evaluation of Figure 7.

The next step in the evaluation procedure consisted of computing objective measurements on the sample images. Information, correlation, mean square error and frequency weighted mean square error measurements were computed.

Rosenfeld [10] briefly considered the problem of determining the information content in an image. To apply information theory to images, he regarded a quantitized digital image as a set of messages by considering the gray level at each picture element as a "message." If there are 2^m gray levels, the total amount of information in an n x n digital image (which is the average amount per element times the number of elements) may be as high as $n^2 \log_2 2^m$ or $n^2 m$ bit. The actual information content depends on the probabilities with which the gray levels occur. The difference between potential and actual information content is called "redundancy."

Again, consider an n x n digital image for which the gray level at each picture level has been quantitized to 2^m values. Clearly, each picture element should be considered as a separate experiment, X. That is, the number of separate experiments is N^2, and a random vector with N^2 elements and its corresponding probability distribution is needed to describe the image. This approach is mathematically untractable; therefore, certain simplifying assumptions must be made.

A great simplification is made if one assumes that the random variables describing each gray level are independent and identically distributed. A good approximation of the

probability distribution for each point may then be made by computing a histogram of the gray level values in a sample image.

An entropy measure, H(X), was computed for each image, X, from probability estimates derived from a 64 gray level histogram for that image. That is, the computation:

$$H(X) = -\sum_{i=0}^{63} p_i \log_2 p_1, \quad \sum_{i=0}^{63} p_i = 1$$

was made for each image. The p_i were computed by counting the number of points in the image which had gray level i and normalizing this number by dividing by the total number of points. The maximum value of H(X), is 6 bits/picture element, and the minimum value is 0.

During the histogram computation, the actual number of distinguishable gray levels was also computed. For example, if the image used every gray level in a 6 bit range, then the image would contain 64 distinguishable levels. The maximum possible number of levels on the University of Missouri image digitalizer is 1024; however, all levels are seldom used because of gain adjustments on the film scanner.

If two images of the same scene are registered in the x and y directions, then a simple measure of how well the images match is easily determined by computing the zero shift correlation between the two images. If this correlation value is normalized by the energies of the two images via the Schwarz inequality then the correlation value of 1 will occur if and only if the image functions are identical to a constant multiple. A correlation value less than 1 indicates that the image functions are not identical.

Andrews [11] has applied the theory of matched filtering to the problem of evaluating the quality of an image with respect to a pre- or postprocessed version of the image. A two-dimensional correlation between images may be made with a matched filter and the resulting correlation value will be a quantitative measure of the “distance” between the two images. The height of a correlation peak at the origin is indicative of the degree of correlation between the two images, and may be used as a measurement for image evaluation, and evaluation decisions can be based on this parameter. A class of matched filters, which includes an energy matched filter and a gradient matched filter, have also been used by Andrews.

A correlation measure, CORR(X,Y), was computed for each image, X, as simply the zero shift normalized correlation between an original scaled image, X, and a processed image, Y. Since the images were registered in the x and y directions, respectively, only the zero shift value was computed. This normalized cross-correlation value is always less than or equal to one, and is only equal to one when the two images are identical.

Mean square error has been used as a fidelity criteria for some image coding systems [12] and is believed not to compare well with subjective human evaluation. The use of frequency weighted mean square error was proposed by Huang [13]; however, no experimental computations were made.

Let f(x,y) represent the input picture brightness function and g(x,y) represent the output picture of an image processing system where f and g are defined for all spatial coordinates (x,y) in a frame or region R.

The error or difference picture is defined as:

$$e(x,y) = f(x,y) - g(x,y)$$

Suppose that the Fourier transform of e(x,y) is given by E(u,v) where (u,v) are spatial frequency coordinates. The mean square error is:

$$D_1(f,g) = \iint_R e^2 (x,y)\, dxdy$$

Using Parsevals' theorem, one may write:

$$D_1(f,g) = \iint_{R_2} |E(u,v)|^2\, dudv$$

where R_2 is the frequency domain of definition of E(u,v).

A weighted mean square error may be defined as:

$$D_2(f,g) = \iint_{R_2} W(u,v)\ |E(u,v)|^2\, dudv$$

where W(u,v) is called the weighting function.

Mean error and mean square error were also computed for the sample images. For each original scaled image, X, and processed image, Y, a difference image, Z = X - Y,. was computed. The mean, p, and the variance, 0, over all picture elements of Z are directly related to the mean square error. The total mean square error, MSE, is given by:

$$MSE = N^2(\mu^2 + \sigma^2)$$

where N^2 is the total number of picture elements.

For comparison, it is easier to use the mean error μ, and the error variance, σ^2, than total mean square error, since μ and σ^2 are normalized to the number of bits/picture element. The frequency weighted function, shown in Figure 9, was used.

For each processed chest X-ray image, entropy, correlation, mean square error, and frequency weighted mean square error were computed. The processed images are shown in Figures 1 through 7, while the calculations are shown in Figure 13. The linearized distribution image was used as the reference image for the correlation and error calculations.

Note that most of the measurements in Figure 13 maintained the distinction between high and low frequency information in the images. For example, the entropy measure, H(X), was much lower for the high pass filtered and high frequency emphasis filtered images than for the original or low pass filtered images. It is also interesting to note that the recursive high frequency emphasis filter had a higher correlation with the reference image than the high emphasis filtered image computed via the FFT. The original scaled image had the largest mean difference because the filtered images were computed from the reference image. The error variance was much higher for the high pass filtered image than for the low pass filtered image, again indicating the difference between low and high frequency information. Also, the frequency weighted error variance was much larger than the unweighted error variance for the low pass filtered and contrast enhanced image, but was only slightly changed for the high pass and frequency emphasized images. Also, note that the correlation measurement follows a shape similar to the radiologists evaluation while the entropy and frequency weighted mean square error follow the same shape as the bioengineers evaluation. Finally, note that the correlation and mean square error measurements are inversely related.

Conclusions Recursive spatial digital filtering and gray level transformation techniques for X-ray image enhancement have been described. These techniques may be implemented with small computer storage requirements and fast computation times and are therefore suited to on-line, small computer applications. The problem of comparing enhanced images was also considered and illustrated by an experimental comparison of subjective and objective evaluations. The comparison illustrated the possible differences between evaluations and the degree of agreement between subjective evaluation and several objective measures of image quality.

REFERENCES

[1] D. Kunio, “Optical Transfer Function of the Focal Spot of X-ray Tubes,” American Journal, Radiation Therapy and Nuclear Medicine, October 1965.

[2] R.A. Morgan, “An Analysis of the Physical Factors Controlling the Diagnostic Quality of Roentgen Images,” American Journal, Radiation Therapy and Nuclear Medicine, December 1965.

[3] E.L. Hall, “Digital Filtering of Images,” Ph.D. Thesis, University of Missouri-Columbia, January 1971.

[4] T.G. Stockham, et. al., “Nonlinear Filtering of Multiplied and Convolved Signals,” IEEE Proceedings, pp. 1264-1291, August 1968.

[5] R. Selzer, “Improving Biomedical Image Quality with Computers,” JPL Technical Report 32-1336, October 1968.

[6] E.L. Hall, R.P. Kruger, S.J. Dwyer, III, D.D. Hall, R.W. McLaren and G.S. Lodwick, “A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images,” IEEE Trans. on Computers (to be published).

[7] J.C. Campbell, “Edge Structure and the Representation of Pictures,” Doctoral Dissertation, Dept. of Electrical Engineering, University of Missouri-Columbia, August 1969.

[8] R.W. Campbell, “The Human Eye as an Optical Filter,” Proc. of the IEEE, Vol. 56, No. 6, pp. 1009-1014, June 1968.

[9] R.H. Morgan, “Threshold Visual Perception and its Relationship to Photon Fluctuation and Sine Wave Response,” American Journal of Roentgenology, Vol. 93, No. 4, pp. 982-997, April 1965.

[10] A. Rosenfeld, “Picture Processing by Computer,” Academic, New York, pp. 7-9, 1969.

[11] H.C. Andrews, “Automatic Interpretation and Classification of Images by Use of the Fourier Domain,” Automatic Interpretation and Classification of Images, Academic, New York, pp. 187-198, 1969.

[12] W.K. Pratt and H.C. Andrews,”Transform Processing and Coding of Images,” University of Southern California EE Report 341, Los Angeles, California, pp. 99-104, March 1969.

[13] T.S. Huang, “Comments on Image Quality,” Massachusetts Institute of Technology Image Processing Short Course Notes, Cambridge, Massachusetts, pp. 1-35:1-37, June 1970.

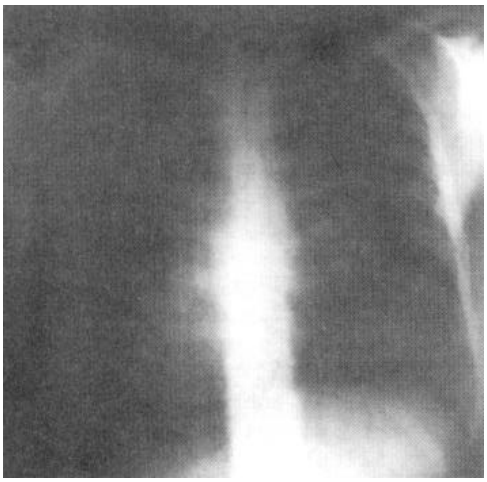

Figure 1. Chest radiograph, scaled.

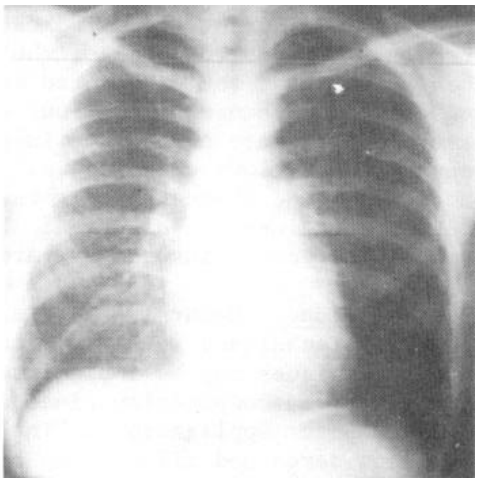

Figure 2. Chest radiograph after distribution linearization.

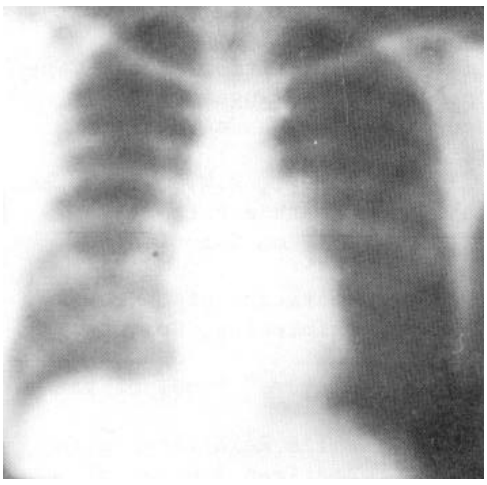

Figure 3. Chest radiograph after low pass filtering

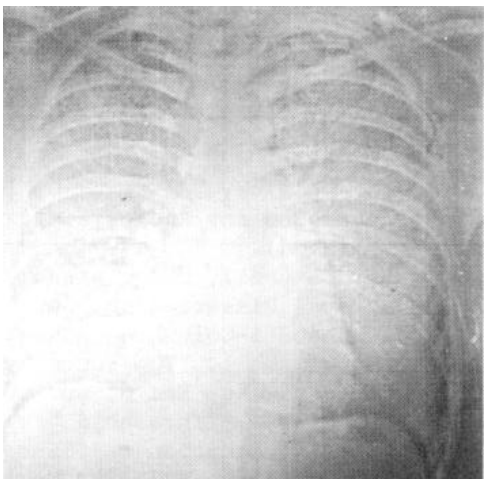

Figure 4. Chest radiograph after high pass filtering

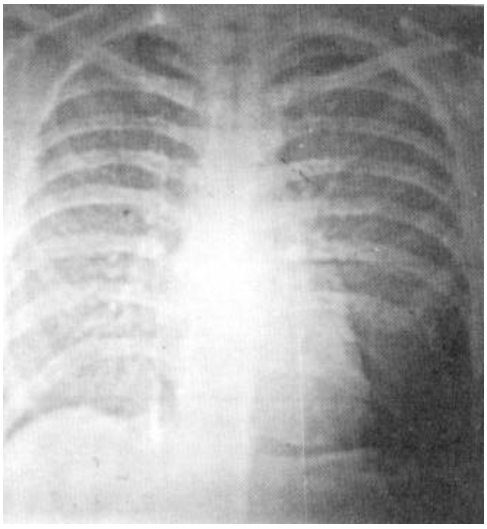

Figure 5. Chest radiograph after high emphasis - FFT.

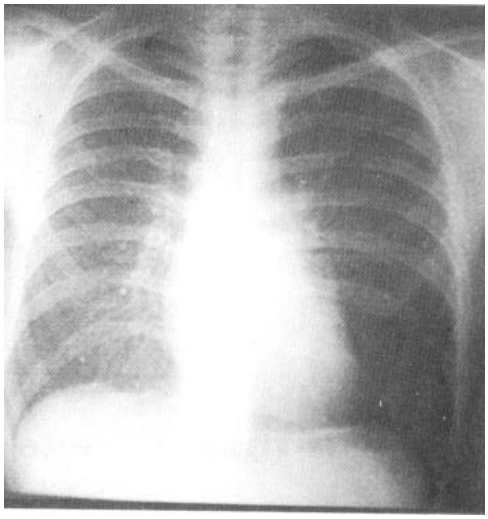

Figure 6. Chest radiograph after high emphasis - recursive filter

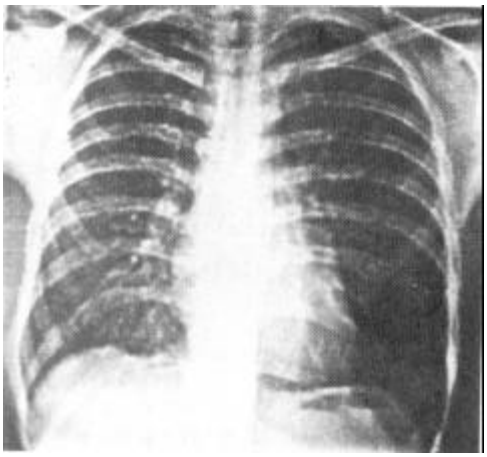

Figure 7. Chest radiograph with high emphasis and contrast enhancement.

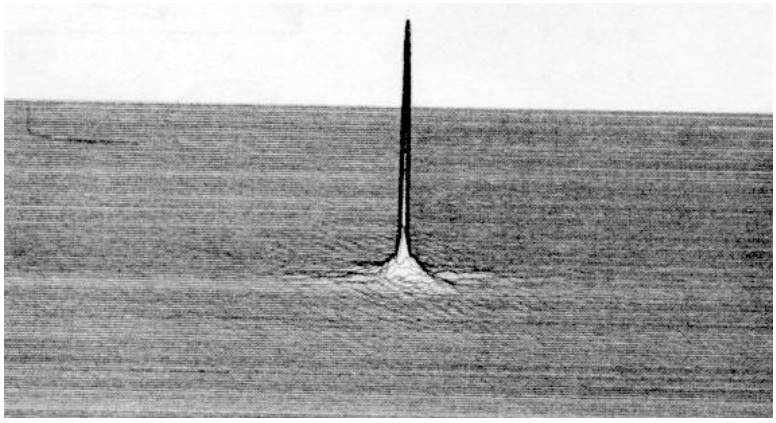

Figure 8. Fourier transform magnitude

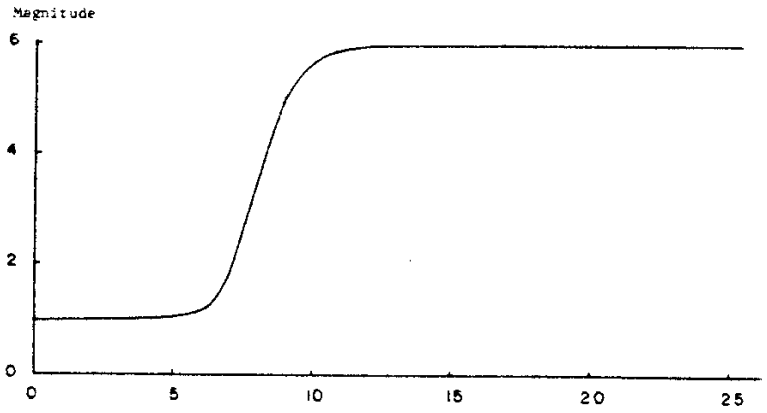

Figure 9. Butterworth high frequency emphasis weighting function with Nyquist frequency at N = 128.

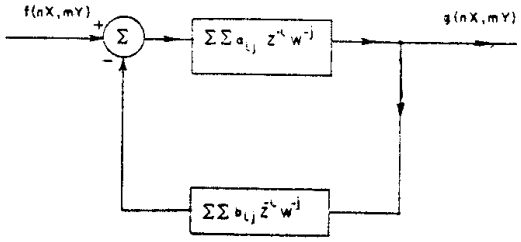

Figure 10. Spatial filtering by recursive partial difference equations.

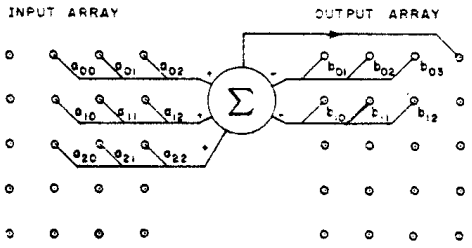

Figure 11. Flow diagram of recursive filter.

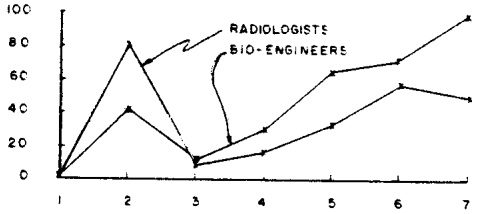

Figure 12. Average subjective goodness functions.

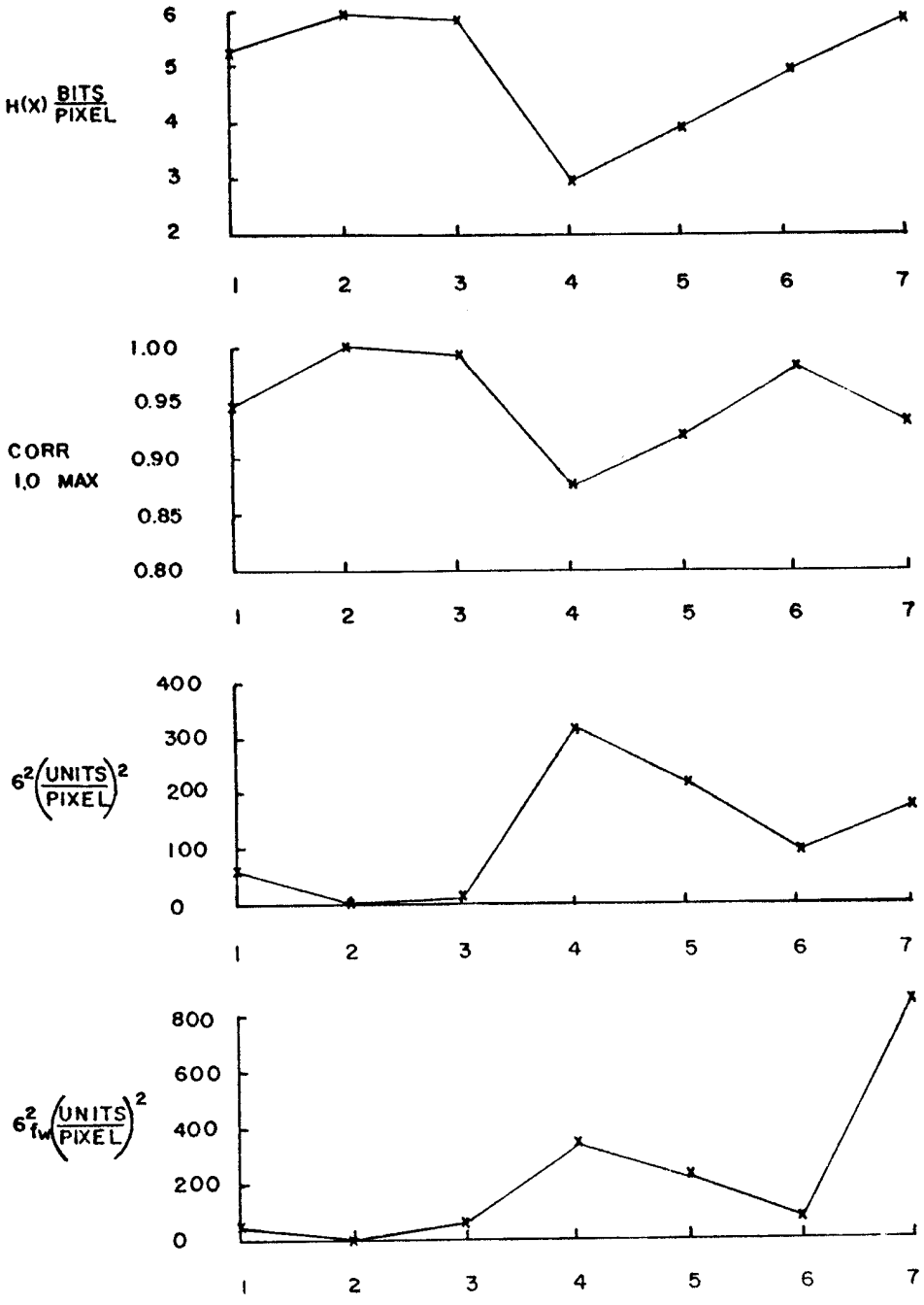

Figure 13. Computed measurements.