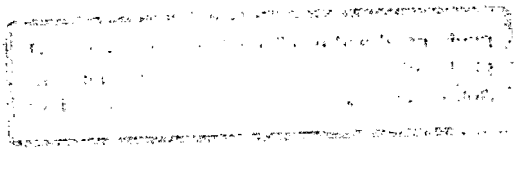
Computer-assisted detection of lung cancer nodules in medical chest x-rays

By Wayne Grant Borchardt

Submitted to the University of Cape Town in fulfillment of the requirements for the degree of Master of Science in Engineering.

Cape Town, September 1992



Declaration

I declare that this dissertation is my own work. It is being submitted for the degree of Master of Science in Engineering at the University of Cape Town. It has not been submitted before for any degree or examination at this or any other university.

Signed by candidate

W. G. Borchardt
(Signature of Candidate)

ate the success rate of all of the enhancement and analysis combinations is required if the automated detection system's success rate is to be maximised. Time and computer resource limitations made such an evaluation impossible.

The experiments which were performed indicated that the automated nodule detection system is capable of detecting camouflaged nodules in digitised chest x-rays. It was, however, also evident that even the best combination of modules results in a poor ratio of true hits to false hits. Increasing the ratio of true hits to false hits can be achieved by improving the operation of the enhancement or analysis module, or by increasing the sophistication of the discriminating mechanisms in the classification process.

Acknowledgements

I would like to thank my supervisor, Professor Gerhard de Jager, for his unfailing enthusiasm and excellent guidance.

I also express my thanks to:

- Peter Moon for his exceptionally insightful contributions at numerous stages in this project.
- De Beers Research Laboratories (DRL) for their generous financial support over the past two years.
- The Foundation for Research Development (FRD) for their financial support over the past two years.
- Fred Hoare for his unselfish dedication in maintaining the computer environment on which I dearly depended.
- Dr Cathy Stoner for her supply and diagnosis of medical x-rays.
- Professor Mike Inggs and Mr John Greene for their technical expertise in the fields of detection statistics and optics, respectively.
- Greg Cox, Alistair Knight, Craig White, Sean Borman, Anthony Kanfer and Trevor Rees for their technical assistance.
- Candice Muller and Christine Riley for their editorial assistance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Diagnostic medicine was revolutionized in 1895 with Röntgen's discovery of x-rays. X-ray photography has played a very prominent role in diagnostics of all kinds since then and continues to do so. It is true that more sophisticated and successful medical imaging systems are available. These include Magnetic Resonance Imaging (MRI), Computerized Tomography (CT) and Positron Emission Tomography (PET). However, the hardware instalment and operation costs of these systems remain considerably higher than x-ray systems. Conventional x-ray photography also has the advantage of producing an image in significantly less time than MRI, CT and PET. X-ray photography is still used extensively, especially in third world countries.

The routine diagnostic tool for chest complaints is the x-ray. Lung cancer may be diagnosed by the identification of a lung cancer nodule in a chest x-ray. The cure of lung cancer depends upon detection and diagnosis at an early stage. Presently the five year survival rate of lung cancer patients is approximately 10%. If lung cancer can be detected when the tumour is still small and localized, the five year survival rate increases to about 40% [12]. However, currently only 20% of lung cancer cases are diagnosed at this early stage [12]. Giger *et al* wrote that "detection and diagnosis of cancerous lung nodules in chest radiographs are among the most important and difficult tasks performed by radiologists".

The observer error which causes nodules to go undetected may be due to [12]:

1. Camouflaging effect of the surrounding anatomical background.
2. Subjective and varying decision criteria used by radiologists.
3. A premature discontinuation of the film reading because of a definite finding.

4. Focusing of attention on another abnormality by virtue of a specific clinical question.

A negative diagnosis occurs when the x-ray is said to be clear of nodules. The diagnosis is false-negative when there actually are nodules but they go undetected. A computer system that alerts the radiologist to the possible occurrence of lung cancer nodules should allow the number of false-negative diagnoses to be reduced. The objective of this thesis was essentially to produce an algorithm which best gives the maximum true-positive[1] and minimum false-negative nodule detection results for lung cancer nodules. In other words a computer-aided nodule detection algorithm of maximum reliability was to be designed with the objective of assisting the radiologist in his diagnosis.

1.1 Definition of the Problem and Motivation for Research

The objective of this research was to locate suspicious nodule sites on a digitized x-ray image. If a technique could be developed that would locate suspicious sites with a fairly high confidence factor, then the cost for a computer aided system would be justified.

A Venn diagram, shown in Figure 1.1, is used to illustrate the principles. $\mathcal{A}$ represents the set of all pixels in the image. $\mathcal{B}$ represents the set of pixels identified as suspicious by the first phase of a computer detection scheme. $\mathcal{C}$ represents the final set of pixels classified by the computer as the centres of nodule sites. $\mathcal{D}$ represents the set of pixels corresponding to the centres of *true* nodule sites. Using set nomenclature for convenience, the following statements are considered as the ideal case:

$\mathcal{B} \cap \mathcal{D} = \mathcal{D}$	set $\mathcal{D}$ is fully included in set $\mathcal{B}$
$\mathcal{C} = \mathcal{D}$	set $\mathcal{C}$ and $\mathcal{D}$ are identical

1.2 A Generic Approach to Automated Detection in Images

The task of the computer, in assisting the radiologist, is essentially automated nodule detection (A.N.D). The *automated* in automated nodule detection implies that the A.N.D system will run without human intervention. The final product of the A.N.D system is an image which has been processed, analysed and classified by the computer. Any suspected nodule

[1] Direct analogy of false-negative, i.e: there are nodules and they are detected.

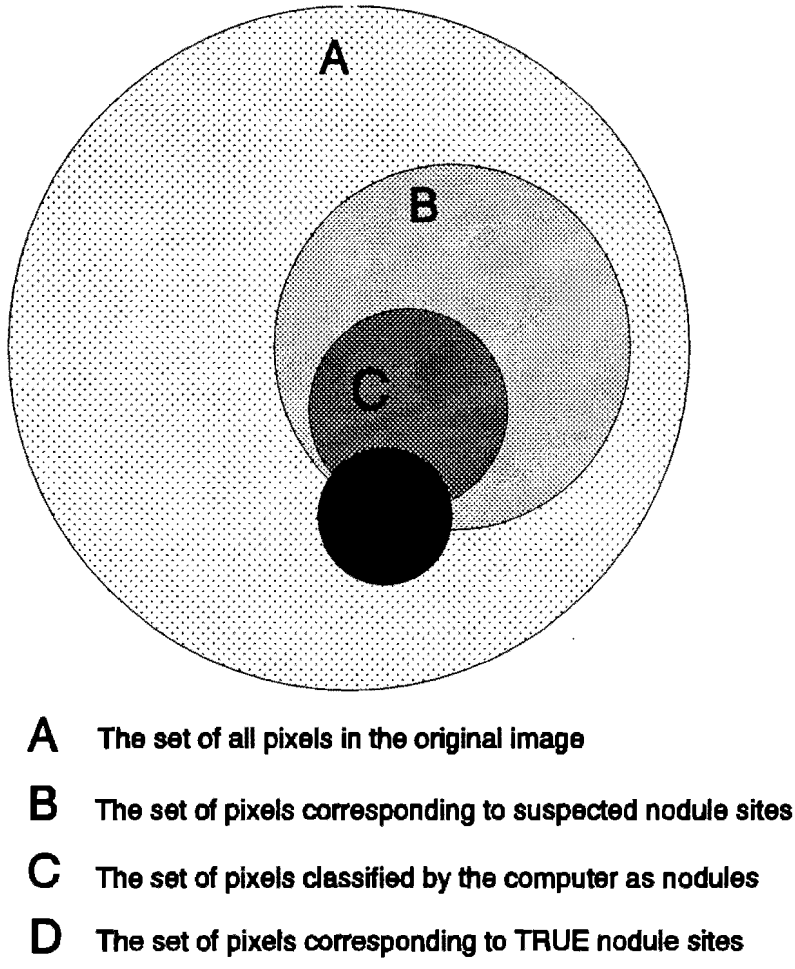


Figure 1.1: Venn diagram illustrating sets of positive and negative nodule sites, as well as those identified by the computer as positive and negative.

sites are highlighted by the computer in the resultant image. A generic breakdown of an A.N.D system is described in this section. A very similar breakdown is presented in [13].

A.N.D may be divided into four primary categories, *viz.* restoration, enhancement, analysis and classification. A block diagram of the generic A.N.D system is provided in Figure 1.2. A module refers to a procedure which may be used as the enhancement or analysis phase of the A.N.D system.

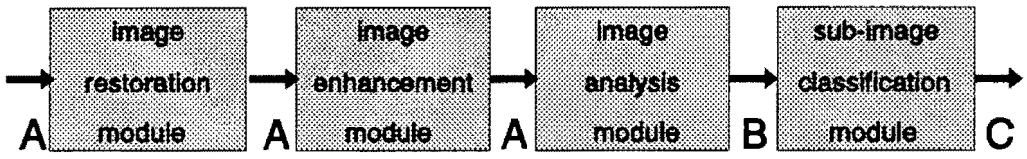


Figure 1.2: Block diagram of the generic Automated Nodule Detection system.

1.2.1 Image restoration

Image restoration, or correction, attempts to remove image characteristics imposed on the image by the nature of the image acquisition. Usually the restoration process attempts to reconstruct or recover an image that has been degraded, by using some *a priori* knowledge of the degradation process [6].

For example, in many optical acquisition systems the restoration process deconvolves the modulation transfer function (MTF) introduced by the lenses. This deconvolution process aims at removing the degradation from the acquired image so that it best matches the original image.

However, in the case of x-ray, the concept of an original image is abstract. The person being x-rayed does not directly convey the original image information. Rather, it is the person's physical x-ray attenuation characteristics which constitute the original image data. The acquired image is thus a function of the original data and the x-ray system's acquisition properties. The acquired image can be improved, or corrected, by removing the artifacts which are imposed on the image by the x-ray system.

1.2.2 Image enhancement

The second stage, in A.N.D, is the enhancement of the image such that it is improved in some sense. In general, the methods of enhancement differ depending on whether the image is for human consideration or if the enhancement is a preprocessing stage - the output of which is intended for further processing by the computer. Sometimes, however, the same enhancement techniques are appropriate for both the human and computer.

The enhancement methods which produce an image for human consumption strive to generate an image which visually best represents the sought information. On the other hand, the en-

hancement for computer consumption produces an image where the information is made most amenable to quantifiable analysis. An automated nodule detection system might therefore use one method of enhancement as a preprocessing stage to further processing, and another enhancement technique to best visually display the final image where suspected nodule sites, set $\mathcal{C}$, have been highlighted.

In general, the enhancement techniques of A.N.D are concerned with simultaneous detail enhancement and dynamic range reduction. Some of the enhancement modules under consideration are unsharp masking, adaptive histogram equalization, statistical difference filtering and morphological processing.

1.2.3 Image analysis

The third phase of A.N.D is the image analysis of the x-ray. Image analysis processes the enhanced image and then outputs abstract data, set $\mathcal{B}$, from which classification may be realized.

The form of the abstract data depends on the type of image analysis used. A common form of the data is a list of the suspected nodule sites, with each entry of the list containing numerous fields. The fields comprise items such as location coordinates and feature parameters of the suspected nodule site.

The image analysis techniques suitable for A.N.D fall into the category of pattern recognition. The methods under investigation include template matching and the Hough transform. Implementations of template matching and the Hough transform are the image analysis modules.

1.2.4 Image classification

The final procedure of A.N.D is classification. The classification process may be likened to the diagnosis made by an expert diagnostician assisted by an intern. The intern points to what he considers possible nodule sites, set $\mathcal{B}$, on the chest x-ray. The expert diagnostician then considers each of the sites indicated by his junior, draws on his own expertise, and makes a decision, generating set $\mathcal{C}$. In the A.N.D system it is the image analysis phase which isolates suspected sites, while the classification phase classifies that site as positive or negative. This classification is based on the information provided in the output list of the analysis phase and perhaps reconsideration of the original image at the locations given in set $\mathcal{B}$. The classification process operates on set $\mathcal{B}$ to produce set $\mathcal{C}$.

1.3 Roadmap to the Dissertation

The structure of this thesis follows the modular nature of the automated nodule detection (A.N.D) system (see Figure 1.2). Chapters successively discuss image enhancement, analysis and classification as modules of the A.N.D algorithm. In general, each chapter has a four part structure, *viz.* (i) an introduction, (ii) the requirements of the module, (iii) an investigation into techniques for the module, and (iv) a summary of the suitability of the techniques examined.

Chapter 2 discusses various classical enhancement techniques, while chapter 3 investigates morphological processing for image enhancement. The enhancement techniques covered in these chapters have two purposes. These are (i) enhancement for improved human intelligibility, and (ii) enhancement as preprocessing for the image analysis phase of A.N.D. The relevant theory for each of these techniques will be discussed, followed by an examination of the technique in terms of the requirements in the A.N.D system. Chapter 4 considers variations on template matching and the Hough transform as potential image analysis modules. In template matching a variance measure, as opposed to correlation measure, is developed. The formation of the x-ray image is discussed when template design is investigated. In the Hough transform section a greyscale implementation is evolved from a binary technique. Chapter 5 discusses the classification of nodule sites from the list of suspected sites produced by the image analysis phase, and describes the use of the receiver operating characteristic in the evaluation of different modules[2] and their respective parameters. Chapter 6 is a collection of the results of most of the above-mentioned techniques applied in practice. Conclusions are drawn and recommendations are made in Chapter 7.

[2]As mentioned in section 1.2, a module is a procedure which may be used as the enhancement or analysis phase of the A.N.D system.

Chapter 2

Classical Techniques for Image Enhancement

2.1 Introduction

Image enhancement is the processing of images to increase their usefulness. When images are enhanced for human viewers, the objective is to improve perceptual aspects. In other applications, an image may be preprocessed to aid machine performance [10]. In the application of the automated nodule detection (A.N.D) system, image enhancement is included for both human consideration and machine preprocessing.

- **Image enhancement for aiding machine performance:** preprocessing the image to aid in subsequent image analysis by the computer.
- **Image enhancement for human consumption:** enhancing the image with the purpose of making it most visually suitable for human interpretation.

The information contained in a medical x-ray has a dynamic range, in general, of 2^{10} to 2^{12}. In other words, for a medical x-ray to be accurately represented as a digital image, the number of intensity quantisation levels should be greater then one thousand. Tests have shown that humans can only discern a maximum of approximately 64, 2^6, grey scales [5]. For the purposes of display it is therefore necessary to reduce the dynamic range, and at the same time ensure that all of the relevant detail is preserved. To achieve dynamic range reduction while preserving the relevant detail requires removing redundancy in the 2^{10} (or 2^{12}) grey levels of the original image.

The objectives of dynamic range reduction and detail enhancement appear to be closely related. If one could remove the slowly varying large scale structure from an image[1], it would appear that the resulting image would contain smaller structures and that the overall dynamic range of the image would be reduced. The net effect should be an image which is easier to interpret. Image enhancement as a preprocessor to machine processing (the image analysis module) also requires that the slowly varying structure is removed from the image and that all the detail is superimposed on a constant level. Figure 2.1 illustrates the point.

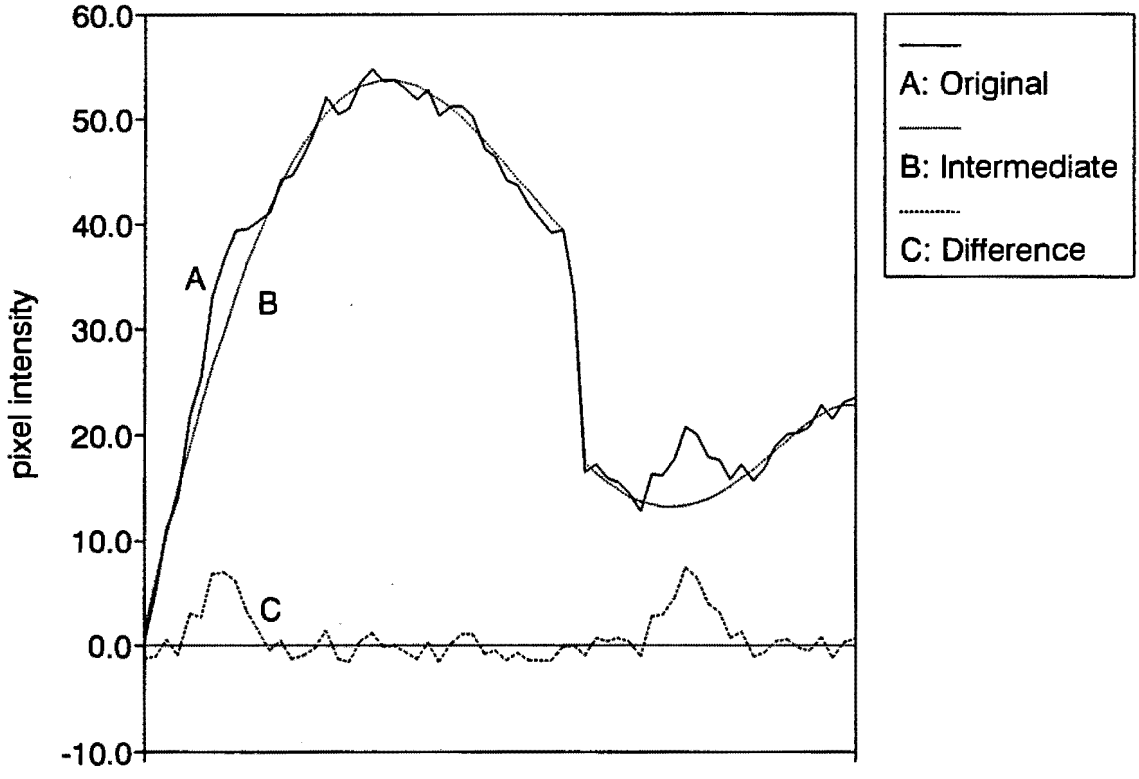


Figure 2.1: Image profile showing dynamic range reduction and detail enhancement using background subtraction.

Although the difference image contains only the detailed structure of the original image it may be added to the original image for the purposes of increasing intelligibility. The resulting image will be a detail enhanced version of the original. Adding the difference image to the original image is optional. In some cases the difference image might itself be a suitably enhanced image. In other cases, though, the difference image contains so little data that it is necessary to add it to the original image for the purposes of putting the detail enhanced

[1]Images are considered to display high intensity for high x-ray attenuation, and vice versa. This is the case for conventional medical x-rays. Images which do not conform to this rule are inverted so that pixel intensity is proportional to x-ray attenuation at that point.

information in context.

It may be necessary to decrease the dynamic range of the resultant image for the purposes of display. The detail-bearing information has already been extracted in the form of the difference image, Figure 2.1C, and therefore the original image, Figure 2.1A, may simply be compressed[2] before adding, to reduce the dynamic range of the resultant image.

In the above discussion, it was assumed that the small scale structures were of low intensity. Tests (and inspection) have shown that this assumption is often invalid. X-ray images contain isolated high intensity small scale structures. The causes of such structures include: dental fillings, steel pins, pacemaker paraphernalia and other metallic components of prosthetics.

There are thus two independent problems in image enhancement for human consumption. They are:

- Enhancing small scale structures
- Reducing dynamic range

The image enhancement requirements for machine and human consumption are described below. Each of the enhancement techniques are evaluated in terms of these requirements.

2.2 Requirements of the Enhancement Module

In order to quantify the merit of each enhancement module it is necessary to define the requirements of the *ideal* enhancement process for both human and machine image enhancement. These requirements were established in an evolutionary manner as more was learned.

The requirements which have been defined for both machine and human needs are:

- **Additive invariance**: enhancement of a nodule should produce an equal output regardless of whether the nodule is superimposed on a flat background or a slope, or a high plateau or a low valley, or superimposed over a discontinuity.

  This idea is illustrated in Figure 2.1 where two similar nodules are superimposed on a background image profile in different local regions. The nodule on the left is superim-

[2]If the original image had a dynamic range of 10 bits and an 8 bit dynamic range is required, this may be obtained by dividing each intensity value in the original image by 4. More generally, N bits are obtained from M bits by dividing the intensity values of the M bit image by 2^{M-N}.

posed on a sloping surface with an average pixel intensity of thirty. The nodule on the right is superimposed on an almost constant background with an average pixel intensity of fifteen. The intermediate profile B is subtracted to yield the difference profile C. The process which produced B can be said to satisfy additive invariance because both of the nodules in C are superimposed on a constant level (zero intensity) and because the nodules have been equally amplified.

Additive invariance is a crucial requirement which will be tested with a simple test image. In the test image two identical nodules are superimposed on a varying background. The first nodule is superimposed on a slope while the second nodule is superimposed on a constant high intensity region. For a process to satisfy additive invariance these nodules must appear identical in the processed image.

The purpose of this additive invariance requirement is to simplify the image for both human and machine consumption. The test image cannot prove the additive invariance of an enhancement technique but it can disprove it.

- **Elimination of structured noise**: it should be ensured that removing the slowly varying structure results in all the detail being superimposed on a constant level.
- **No artifact production**: artifacts are undesirable byproducts of a process and can be considered noise.
- **Rotational invariance**: this requirement ensures that an object which is circularly symmetric before processing will remain circularly symmetric. Another consequence of rotational invariance is that a processed rotated object will be identical to a processed non-rotated object, albeit rotated.
- **Robust operation**: the enhancement operation should be insensitive to noise.
- **Short processing time and cheap installation costs**: if an A.N.D system is to be implemented as a screening technique, the processing time should be short and the installation costs of the system should be small. Both of these items are a function of the computational requirements of the system.

As mentioned in section 2.1 additional requirements for human consumption have been identified:

- **Reduced Dynamic Range**: the number of bits required to store all the information is best reduced to a maximum of eight. This is justified as follows: (i) As stated in section 2.1, human perception of greyscale images is limited to six bits. (ii) Commercial eight bit display devices are available and affordable. This hardware therefore copes adequately with the human requirements for a greyscale display.

- **Detail Enhancement**: It is important, for human consumption, that reducing the dynamic range is done in such a way that the relevant detail is made more visible. This amounts to amplifying the detail with respect to the background. In effect, detail enhancement is high pass filtering. One can, however, achieve vastly different results by using variations on the theme of high pass filtering. For instance, high pass filtering may be linear or nonlinear, adaptive or non-adaptive, or even adaptive nonlinear.

2.3 Image Enhancement Modules

In this section all of the classical enhancement techniques which were attempted in this research will be discussed. Morphological image enhancement techniques are the subject of chapter 3. Many of the techniques attempt to generate an intermediate image which contains the underlying background image. This intermediate image is then subtracted from the original image to yield an enhanced image.

In the case of nodule detection in chest x-rays, the background image is the healthy chest x-ray and the lung cancer nodules are the image detail of interest. The background information may be thought of as structured noise. The complexity of the problem arises because of the complexity of the structured noise. An x-ray image, like an optical image, may contain incredible diversity, including discontinuities. These discontinuities are the primary cause of the shortcomings of linear processing techniques.

2.3.1 Unsharp masking

Unsharp masking involves subtracting a low pass version of an image from the original image. This is essentially a high pass filtering technique inherited from photography. In the photographic process, a film is exposed through a negative superimposed on a slightly defocused positive transparency. In this way, the local mean is subtracted and the result is an image with improved edges [20, pp84-85].

Unsharp masking is a popular digital image processing technique which enhances local contrast and improves the aesthetic appeal of the image. The output image Y is generated according to the formula:

$$Y = cS - (1 - c)S_{LP}$$

where,

c is a gain parameter

S is the original image
S_{LP} is a low pass version of S

Linear unsharp masking

Conventionally, the low pass image is obtained using linear low pass filtering. S_{LP} may, however, be produced by any variation of low pass filtering, including adaptive and/or nonlinear techniques. Regardless of the type of low pass filtering, S_{LP} is always the background image which is subtracted from the original image to produce the enhanced image.

Linear low pass filtering may be performed in either the space domain or Fourier (spatial frequency) domain. The Fourier domain is often used to lessen the computational burden by making use of the fast Fourier transform (FFT). Regardless of the means of implementation of low pass filtering, comprehension of linear filtering in the Fourier domain enhances understanding of the filter in the space domain.

For instance, the dimension of the local region (footprint)[3] of averaging for the low pass filter has a marked effect on the resultant image. In particular, using a large footprint in the space domain amounts to a narrow spatial frequency band low pass filter. Such a filter maintains only the very low spatial frequency components, corresponding to the slowest varying structures, in the low passed image S_{LP}. Consequently it is only these very low spatial frequencies which are attenuated in the difference image Y. At the other extreme, a small footprint (for example, averaging a local 3x3 region) has a broad response in the spatial frequency domain. This has the effect of attenuating only very high spatial frequencies in S_{LP}, resulting in these high frequency components existing in Y. With a small footprint Y appears as an edge detected image because it is only the edges and very small structures which contain these high spatial frequency components.

The shape of the footprint also influences the effect of the low pass filter. This too can be understood by inspecting the spatial frequency response of the filter. Computational requirements of linear low pass filtering can be significantly improved with the use of a separable[4] filter. The shape of the filter determines if it is separable. For example, a square footprint is separable while a circular footprint is not. Although a separable filter is desired for computational reasons, the spatial frequency response of most separable filters is inferior. Clearly a

[3]The terms local region, neighbourhood, local neighborhood, contextual region, footprint and window are all used interchangeably. In most cases the terms used by the referenced authors are adopted.

[4]A two dimensional filter $f(x, y)$ is separable if f can be written as $f_1(x)f_2(y)$.

tradeoff must be made when designing the most suitable linear filter for an application.

A further factor in the design of the low pass filter is the windowing function. Window functions are employed to lessen the sidelobe effect of the filter response in the transform domain.

Median unsharp masking

Median filtering is an example of *non*linear low pass filtering. Because it is nonlinear, median filtering cannot be performed in the Fourier domain. Median filters may be implemented in a variety of ways by modifying the shape and size of the footprint. For instance, the footprint from which the median is derived may be square, circular, ring or cross shaped. A square footprint may also have the median computed for each of its rows, and then the median taken of that set of medians. One dimensional (1D) median filtering may be used horizontally on the image to produce an intermediate image. The intermediate image is then median filtered using 1D vertical median filtering to produce the background image[5].

Examination of unsharp masking

Artifacts are generated if the image is low pass filtered using a simple linear averaging process (see Figure 2.2). It is interesting to note that it is exactly these artifacts that provide the aesthetic appeal of the image produced by conventional unsharp masking. The artifacts appear as shadow bands on either side of an edge. The intensity of the artifacts is proportional to the steepness of the slope, and the width of the artifact is proportional to the dimension of the low pass filter.

The linear unsharp masking process appears to satisfy the requirement of additive invariance when applied to the test image (see Figure 2.3B). Figure 2.3B was generated using a square footprint of dimension 11x11 pixels and $c = \frac{1}{2}$. The generation of an artifact is also evident in this figure.

Use of a two dimensional (square, ring or circular footprint) median filter as the low pass filter produces substantially better results. It does, however, also generate artifacts at corners, as illustrated by Figure 2.4 where a square footprint 17x17 was used on a 2D unit step function. A cross footprint median filter does not produce these artifacts if, and only if, it is aligned[6]

[5]The process could also be run vertically first, and then horizontally. Note that this will produce a slightly different image.

[6]Aligned in this sense implies that the *limbs* of the footprint are at 0° and 90° to the edges of the corner. If

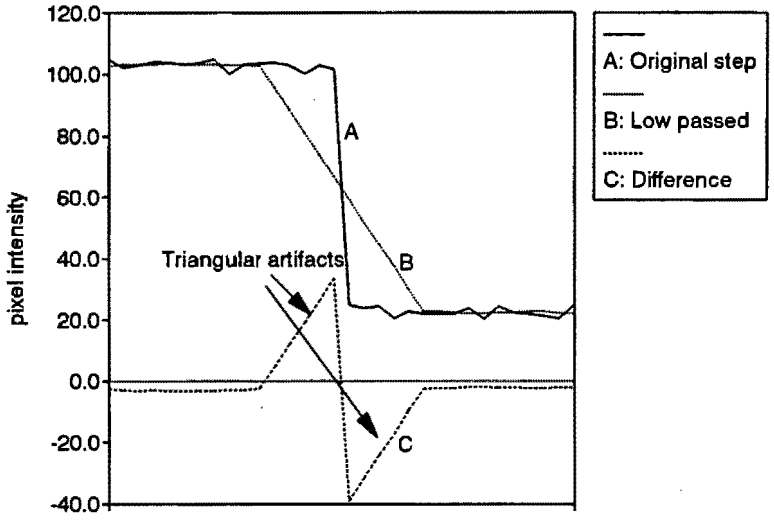


Figure 2.2: The generation of artifacts using the averaging of linear unsharp masking.

with the corner. The effect of taking the median of the set of medians of each row produces very similar results to a simple square median filter.

The corner artifacts, resulting from all 2D median filters, may be eliminated by operating the median filter as two one dimensional operations. It should be noted that the median filter is not separable, and consequently these one dimensional operations produce a different result from the two dimensional filter.

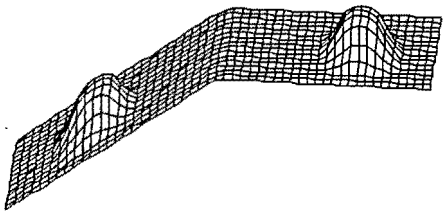
The square footprint median unsharp masking process appears to satisfy the requirement of additive invariance when applied to the test image (see Figure 2.3C). Figure 2.3C was generated using a 17x17 square window and $c = \frac{1}{2}$.

2.3.2 Statistical difference filtering

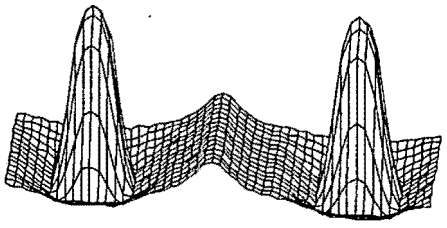
The statistical difference (SD) filter was evaluated by Morishita and Akimbo [19]. They required a preprocessing filter which would satisfy the following requirements:

- emphasize the fine structure of the image
- eliminate the effect of slow grey level fluctuations

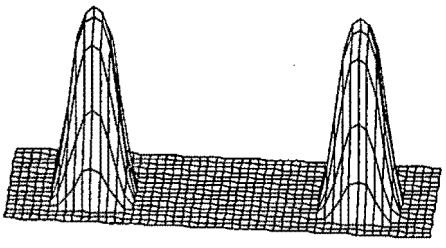
this is not the case, for instance if the limbs are at 45° to the corner, then similar artifacts to the symmetrical median filter result.



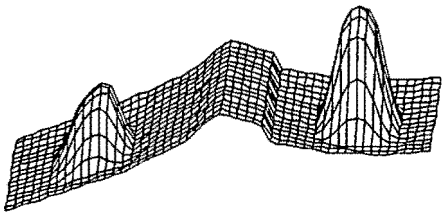
A. Original Test Image



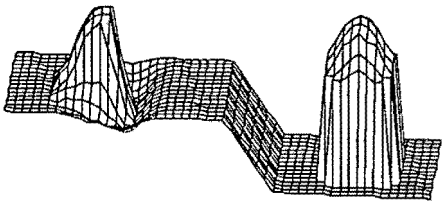
B. Unsharp Masking



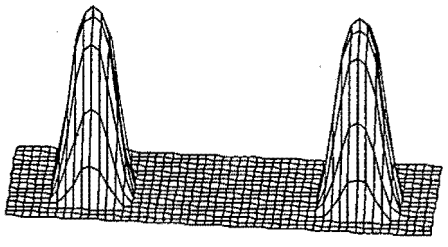
C. Median Unsharp Masking



D. Statistical Difference Filtering

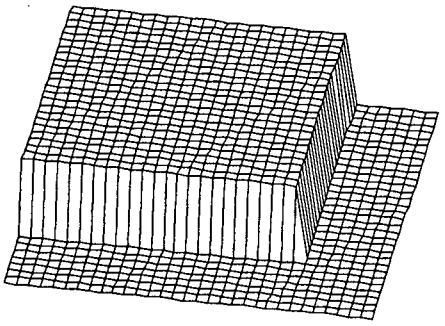


E. Adaptive Histogram Equalisation

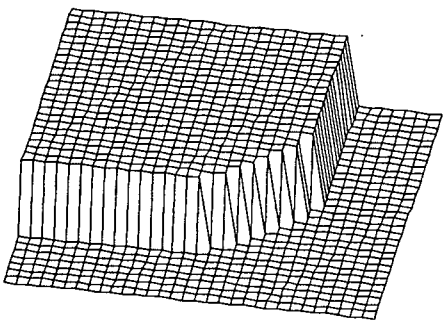


F. Morphological Greyscale Opening

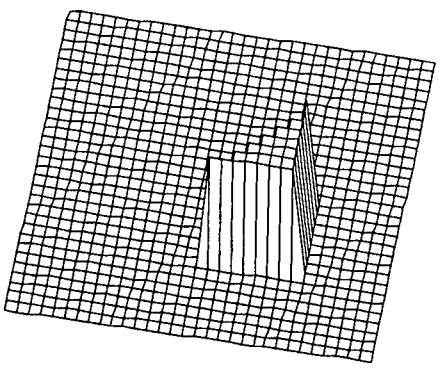
Figure 2.3: Test for additive invariance using different enhancement techniques.



A. Original Test Image



B. Median filtered image



C. Difference image C = A - B

Figure 2.4: The generation of the corner artifact from the symmetrical 2D median filtering process.

- be invariant to the properties of the image
- short processing time

The first two requirements are essentially detail enhancement and dynamic range reduction. The third point specifies that the filter should be suitable for any image. The final point is clear. These requirements, perhaps with the exception of point three, are a subset of the requirements outlined in section 2.2.

Morishita and Akimbo claim that the SD filter outperforms the other filters which they evaluated. These filters included the differential filter (implemented as a Prewitt operator), the Laplacian, a bandpass filter, and a highpass filter.

The SD filter is implemented in a pixel by pixel, adaptive processing manner. Every pixel in the output image is evaluated according to:

$$Y_{ij} = \alpha(S_{ij} - \mu_{ij})\frac{\overline{\sigma}}{\sigma_{ij}} + \mu_{ij}$$

where,

Y_{ij} is the intensity of the output pixel
α is a gain parameter
$\overline{\sigma}$ is the standard deviation for entire image
σ_{ij} is the standard deviation for the local region
μ_{ij} is the mean value of the local region
S_{ij} is the intensity of the current pixel in the input image

Examination of the SD filter

The SD filter is difficult to evaluate because it requires two user input parameters, *viz.* gain factor α and local region size. Optimising this filter requires an exhaustive search of a two dimensional (2D) space. The 2D search space results from the two user input parameters. The parameters to optimise the filter have not been established because the filter has been shown to be inappropriate. Firstly, the filter generates artifacts, and secondly, the filter is *not* additively invariant. This lack of additive invariance is illustrated in Figure 2.3D.

As with unsharp masking, the artifacts improve the aesthetic appeal of an image produced by SD filtering. The artifacts appear as shadow bands where there are sharp differences in image intensity.

The SD filter enhances small scale structures with some success. With a small footprint the artifacts are negligible and the detail enhancement is still noticeable. The SD filter is thus a worthwhile means of improving the aesthetic appeal of an image.

2.3.3 Adaptive histogram equalization

The process of histogram equalisation attempts to transform the image histogram into a uniform histogram. This is achieved by computing a mapping function, or look-up table, between the cumulative distribution functions of the image histogram and the desired uniform histogram. A comprehensive explanation of histogram modifications, like equalisation, is

found in [17]. Histogram equalisation is a fast and simple means of contrast enhancement. Due to its global application, it is often ineffective at accomplishing local contrast enhancement. The technique discussed in this section, adaptive histogram equalisation (AHE), adaptively performs histogram equalisation on local regions within the image. This has the advantage of achieving local contrast enhancement.

AHE is a contrast enhancement technique in which the intensity of each pixel is replaced by its rank within a local neighbourhood [11]. Each pixel in the output image is computed according to the AHE algorithm [11]:

```
for each pixel (i, j) in image do
begin
    rank = 0
    for each (x, y) in contextual region of (i, j) do
    begin
        if image[i, j] > image[x, y]
            then rank = rank + 1
    end
    output[i, j] = rank
end
```

Contrast limited AHE

As an undesirable consequence of enhancing image contrast, AHE tends to enhance noise in relatively constant level areas of the image [11]. Contrast Limited AHE (CLAHE) attempts to lessen the enhancement of noise by varying the maximum possible level of contrast enhancement according to the shape of the histogram in each contextual region.

Examination of AHE

CLAHE was employed in [11] to lessen the over-enhancement of noise. CLAHE requires two user input parameters, *viz.* clip level and window size. This makes it difficult to evaluate for the same reasons as the statistical difference filter.

Rehm and Dallas revealed a further disadvantage of AHE and CLAHE [23]. This disadvantage is the generation of a zone of *reduced* contrast at regions where a substantial change in intensity occurs. This artifact is due to the large changes in the local histograms as they cross a boundary. The width of this artifact is directly proportional to the dimension of the

contextual region.

Rehm and Dallas devised a scheme for suppressing the boundary artifact. It involved preprocessing the image with background subtraction before subjecting the image to CLAHE. Their background subtraction algorithm used interpolation of cubic splines between knot points. The knot spacing was determined empirically and had not yet been automated.

Operation of AHE on the test image clearly indicated that AHE is *not* additively invariant (see Figure 2.3E).

2.3.4 Nonlinear adaptive curve fitting

Nonlinear adaptive curve fitting attempts to generate the underlying background image without being influenced by the small scale structures which are superimposed on this background. It operates on the assumption that the slope of the one dimensional profile before and after a nodule is approximately equal.

The algorithm operates in a one dimensional manner. Each row of the image is processed to produce an intermediate image. This intermediate image is then processed columnwise[7] by the same algorithm.

The operation of one dimensional nonlinear adaptive curve fitting can be understood by considering a cursor c_a which follows the given profile and another cursor c_b which traces out the new profile. c_a moves along the original profile, computing the instantaneous slope of the profile. While the slope remains constant, within a given tolerance, c_b mimicks c_a. If the slope deviates sufficiently then c_a continues alone for a given distance. If, within this distance, the slope returns to its value before the deviation, then it is assumed that a small scale structure was encountered. In this case c_b generates the new profile, using the slope value before the deviation. In this way an underlying profile, which does not contain small scale structures, has been created. A difference profile would show up all the small scale structures.

The algorithm for performing nonlinear adaptive curve fitting, in one dimension, is provided below. W is the maximum sought after width, X is the array of the original data set, and Y is the array of the modified data set.

[7] Alternatively the image may be processed columnwise and then rowwise. Note that this will produce a slightly different image.

```
for every pixel i in a row do
begin
    1. oldslope = Calculate Slope( X )
    2. oldvalue = X[i]
    3. move to next data point
    4. currentslope = Calculate Slope( X )
    5. is oldslope ≈ currentslope?
    6. if yes then
    begin
        a. Y[i] = X[i]
        b. oldslope = currentslope
        c. goto 3
    end
    else
    begin
        position = 0
        move ahead one point at a time, computing the current slope, until
            (i) (currentslope ≈ oldslope) AND (currentvalue ≈ oldslope.position +
            oldvalue) OR (ii) we have moved W data points ahead
        increment position
    end
    7. if (i) then (a small scale structure was encountered, generate the underlying slope)
    begin
        a. return to original current position and generate Y using oldslope for as many data
        points as were used.
        b. compute currentslope at the final position.
        c. oldslope = currentslope
        d. goto 3
    end
    else (ii) (no small scale structure encountered, an actual slope change in the underlying
    profile occurred)
    begin
        a. return to original current position and generate Y as Y[i] = X[i] for those W data
        points.
        b. compute currentslope at final position
        c. oldslope = currentslope
        d. goto 3
    end
end.
```

Examination of nonlinear adaptive curve fitting

This approach suffers from several flaws. Firstly it assumes that the small scale structures are superimposed on linear profiles. This is a gross oversimplification. The second problem arises when the *approximately equal to* conditions are evaluated. What tolerance must be allowed to ensure values may be decided on as being approximately equal?

The evaluation of additive invariance for nonlinear adaptive curve fitting is difficult because of the tolerance problem discussed above. If the slope on which a nodule is superimposed is sufficiently steep, then additive invariance will fail. This is because the change in slope due to the nodule is not adequate to exceed the tolerance of what constitutes a constant slope. Also, if a nodule is superimposed on a slope which is not linear, the intermediate profile will be linear (by the nature of the process) and the resultant nodule will be distorted. In other cases, however, it is likely that additive invariance will be satisfied.

2.4 Summary of the Classical Enhancement Modules

The classical enhancement modules are summarized in terms of the requirements described in section 2.2. A table for the evaluation of additive invariance, structured noise elimination, artifact production, rotational invariance, robust operation and processing time is provided (see Table 2.1).

Processing time is written in order notation, followed by a label which represents the predominant[8] type of operation. All of the algorithms process each pixel and therefore the order notation is provided as a comparison basis on the computational requirements for a single pixel. Thus, the processing time for an entire image is proportional to the product of the image dimensions and the computational cost per pixel. In most cases the order notation is, by necessity, an approximation. The labels are (i) **boolean**, representing boolean comparison, (ii) **integer**, integer operations, and (iii) **floating**, for floating point operations.

[8]If floating point operations are used at all, it is likely that the **floating** label will be used. In other cases the operation which is used most in the algorithm is written as the label.

[9]Yes, when symmetrical two-dimensional median filters are used, but no when two one-dimensional median filter passes are used.

[10]This process is exceedingly slow when sorting is used but may be sped up significantly with the use of a histogram method. The median may be found from a data set by generating a histogram of that data set and then counting the bins of the histogram until half the total count is reached. The bin which results in the count exceeding half the total count corresponds to the intensity value which is the median of that data set.

Enhancement modules	Additive invariance	Removal of structured noise	Artifact production	Rotational invariance	Robust operation	Processing time
Linear Unsharp Masking	YES	YES - uses background subtraction	YES - along edges	YES	NO	$O(n^2)$ integer
Median Unsharp Masking	YES	YES - uses background subtraction	YES - at corners[9]	YES	NO	$O(n^3)$[10] integer
Statistical Difference Filtering	NO	to some extent	YES - along edges	YES	NO	$O(n^2)$ floating
Adaptive Histogram Equalisation	NO	to some extent	YES - along edges	YES	very poor	$O(n^2)$ boolean
Nonlinear Adaptive Curve Fitting	in most cases	YES - uses background subtraction	unknown	NO	NO	$O(n)$ floating

Table 2.1: Comparison of classical enhancement techniques.

With regard to enhancement for best visual representation, reduced dynamic range and detail enhancement are also taken under consideration. All the modules which use background subtraction will provide dynamic range reduction for most of the image. Isolated regions of high intensity might still exceed the required dynamic range. Dynamic range reduction in statistical difference filtering and adaptive histogram equalisation is difficult to evaluate. However, neither of these methods exhibit additive invariance and are thus withdrawn as candidates for the module of image enhancement.

Nonlinear adaptive curve fitting is also not guaranteed to satisfy additive invariance. It, like one dimensional median filtering, will produce different outputs depending on whether the image is processed vertically or horizontally first. Nonlinear adaptive curve fitting has additional problems, as mentioned in section 2.3.4, and will not be considered further.

All the enhancement modules perform detail enhancement. Disregarding statistical difference filtering, adaptive histogram equalisation and nonlinear adaptive curve fitting, we are left to decide between linear unsharp masking and median unsharp masking.

Chapter 3

Morphological Processing for Image Enhancement

3.1 Introduction

This chapter discusses the use of morphological image processing techniques for image enhancement in three main sections. In the first section, the relevant morphological theory is presented. The second section investigates the use of a new algorithm, namely the rotating rod algorithm. The third section discusses the use of a fundamental morphological operation, the top-hat transform, which was found to be the most appropriate. Both the rotating rod and the greyscale top-hat transform are examined as potential candidates for the image enhancement module. As such, they are evaluated according to the requirements presented in section 2.2.

3.2 Morphological Image Processing Theory

The content of this section is based primarily on the content of two papers [18, 27] and is intended as a summary of the morphological image processing theory relevant to this thesis. A more complete mathematical description of morphological processing is provided in [18, 27, 14].

3.2.1 Introduction to morphology

Classical techniques (theory of linear systems and Fourier analysis) are often of limited use because they do not directly address the fundamental issues of how to quantify "shape or geometrical structure" in images [18]. In contrast, mathematical morphology can rigorously quantify many aspects of the geometrical structure of a signal, in a way that agrees with human intuition and perception [18].

An interesting perspective on morphological processing is presented by Serra [25] where he proposes "... our philosophy will consist of stating that the images under study exhibit too much information, and that the goal of any morphological treatment is to *manage the loss of information* through the successive transformations."

3.2.2 History and development of mathematical morphology

Morphological image enhancement is a nonlinear approach to image enhancement based on the theory of mathematical morphology. The original formulation of mathematical morphology was proposed in the mid 1970's and was limited to binary images. The concepts were then extended in the mid 1980's to account for greyscale images [18].

Extension from binary to greyscale morphology may be achieved in various ways [18]. Traditionally Serra used the representation of a n dimensional (n-dim) function[1] $f(x)$, (x is a n-dim vector, $f(x)$ is a greyscale signal[2]) by the ensemble of its threshold sets. An alternative extension by Sternberg represents a n-dim function $f(x)$ by a (n+1)-dim set, its umbra. The umbra of f is $\mathcal{U}(f) = \{(x, a) : a \leq f(x)\}$. It is this extension, using umbrae, which will be discussed. An illustration of the umbra of a simple test image is provided in Figure 3.1.

In [27] the equivalence between morphology in greyscale images and the more established field of mathematical morphology is described. Sternberg's [27] explanation of greyscale morphology is essentially an explanation of mathematical morphology followed by reasoning which indicates that greyscale images are, in fact, the same mathematical sets for which mathematical morphology is applicable. Consequently the theory of mathematical morphology is directly relevant to greyscale morphology, which is necessarily an application of mathematical morphology. Sternberg's explanation is summarized below:

[1] In the case of an image f is a 2 dimensional function, n=2.

[2] f has a continuous or a quantized range.

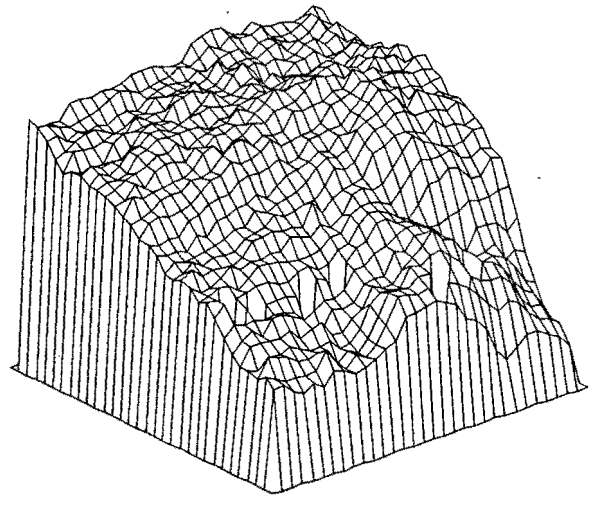


Figure 3.1: The umbra of a simple test image is the solid set which extends infinitely below the undulating image surface.

1. Consider morphology of sets in Euclidean 3D space. Of particular interest are the umbrae. These are solid sets which extend unbroken indefinitely downward in the negative z direction.

2. The details of an umbra's third dimension can be determined by a single parameter, the height z of the umbra at (x, y) in the image domain.

3. Umbra means shadow, and the umbra of a set $\mathcal{X}$ in 3 dimensional space includes both $\mathcal{X}$ and the volume of points in its shadow (the shadow is cast by a point light source at an infinite distance in the positive z direction).

4. A greylevel function f can be thought of in Euclidean space as a set of points $[x, y, f(x, y)]$, imagined as a thin undulating, not necessarily connected sheet. A greyscale image $f(x, y)$ is represented in mathematical morphology by an umbra $\mathcal{U}[f]$ in Euclidean 3D space where a point $p = (x, y, z)$ belongs to the umbra if, and only if, $z \leq f(x, y)$.

3.2.3 Greyscale morphological transformations

Umbrae remain umbrae under the usual morphological transformations of union and intersection, dilation and erosion (to be discussed). Note that union is the term used in mathematical morphology (and also in set theory), and that dilation is used in greyscale morphology. The same is true for intersection and erosion. The equivalence of greyscale images and umbrae allow us to speak morphologically of dilating or eroding a greyscale image by a structuring

element[3]. An explanation of dilation and erosion follows.

Dilation and erosion

Dilation and erosion are fundamental operations of morphology, both binary and greyscale. Composite operations like openings and closings (to be discussed) are expressible in terms of dilations and erosions only. Dilation and erosion are carried over from set theoretic operations called *Minkowski addition* and *Minkowski subtraction*, respectively [14]. A comprehensive study of dilations and erosions is found in [14].

The dilation of an umbra $\mathcal{U}[\mathcal{X}]$ by a structuring element (SE) $\mathcal{B}$ is formally stated as the maximum of the paired summations of $\mathcal{B}(x,y)$ with each of the points of $\mathcal{X}(x,y)$; where $\mathcal{U}[\mathcal{X}]$ is the umbra of image $\mathcal{X}$ and $\mathcal{B}(x,y)$ is the function describing the surface of $\mathcal{B}$. The dilated image D is then

$$D(x,y) = \max_{i,j}[\mathcal{X}(x-i, y-j) + \mathcal{B}(i,j)].$$

Using morphology nomenclature, we have $D = \mathcal{X} \oplus \mathcal{B}$. Similarly, the erosion of $\mathcal{U}[\mathcal{X}]$ by $\mathcal{B}$ is the minimum of the paired differences of $\mathcal{B}(x,y)$ with each of the points of $\mathcal{X}(x,y)$. The eroded image E is then

$$E(x,y) = \min_{i,j}[\mathcal{X}(x-i, y-j) - \mathcal{B}(-i,-j)].$$

Using morphology nomenclature, we have $E = \mathcal{X} \ominus \mathcal{B}$. Note that the negative arguments of $\mathcal{B}$, i.e: $\mathcal{B}(-i,-j)$, are carried over from the formal mathematical definition. In all instances of practical SE's we have symmetry, thus $\mathcal{B}(i,j) = \mathcal{B}(-i,-j)$.

Sternberg [27] makes the observation that the expressions for dilation and erosion are similar to the convolution integral, or convolution sum in the discrete case. Unlike linear convolution, however, dilations and erosions are non-invertible [27]. Sternberg thus concludes that "Image Processing through iterative morphological transformations is a process of selective information removal where irrelevant image content is irrecoverably destroyed, enhancing the contrast of essential image feature" [27]. This is notably similar to Serra's quote in the introductory section.

[3]The geometric description of a morphological structuring element is analogous to the parametric description of a linear filter. Structuring elements will be discussed in more depth later.

Opening and closing

Openings and closings are composite operations of dilations and erosions. An opening of $\mathcal{X}$ by $\mathcal{B}$ is denoted as $\mathcal{X} \circ \mathcal{B} = (\mathcal{X} \ominus \mathcal{B}) \oplus \mathcal{B}$. A closing of $\mathcal{X}$ by $\mathcal{B}$ is denoted as $\mathcal{X} \bullet \mathcal{B} = (\mathcal{X} \oplus \mathcal{B}) \ominus \mathcal{B}$.

Openings, closings and the fundamental operations of erosion and dilation are conceptually described in the next section. Thus far the four morphological operations, dilation, erosion, opening and closing, have been discussed. Although these operations have been extended for manipulation of greyscale images, their operation is simpler to visualise when applied to binary images. A graphical depiction of the fundamental morphological operations applied to a binary image is provided in Figure 3.2, taken from [18].

3.2.4 Conceptual description of morphological operators

In this discussion a conceptual, as opposed to mathematical, description of morphological processing is presented. In simple terms, morphological processing is achieved by moving a structuring element (SE), of some geometrical shape, above or below the image surface.

A conceptual[4] description of erosion (and dilation) is provided because it helps to clarify the otherwise abstract mathematical descriptions most often found in the literature. The description will be given in terms of erosion but it may be read in terms of dilation by replacing the angle bracketed word with the preceding word. And, by extension, opening and closing are also described because an opening and a closing are compound operations of dilation and erosion. This algorithm describes the generation of the eroded <dilated> image surface S' from the original image surface S.

[4]The conceptual algorithm may differ considerably from the implementation algorithm. The conceptual algorithm is provided here because it is easier to understand. The implementation algorithm is often dependent on the structuring element if the algorithm is optimised.

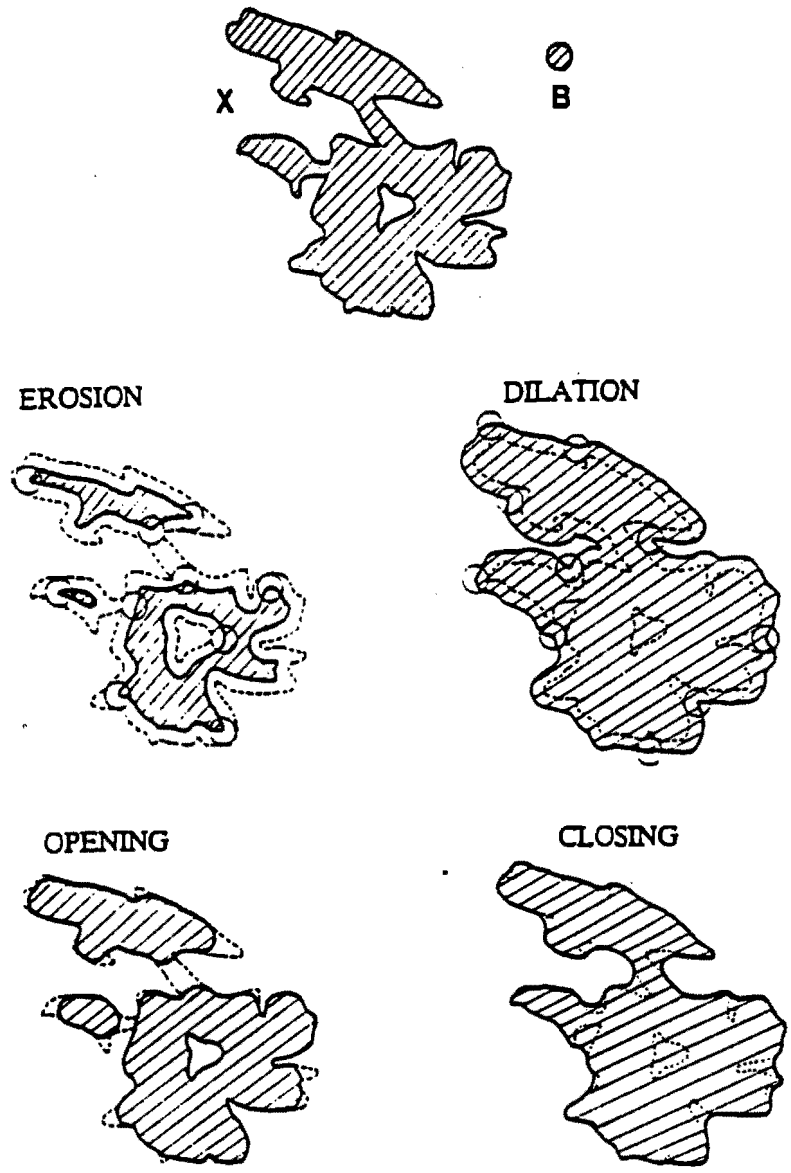


Figure 3.2: Erosion, dilation, opening and closing of a binary image $\mathcal{X}$ by a structuring element disk $\mathcal{B}$. The shaded areas correspond to the interior of the sets, the dark solid curve to the boundary of the transformed sets, and the dashed curve to the boundary of the original set $\mathcal{X}$. Taken from [18].

For every pixel (i, j) in S do

1. Position the origin of the structuring element (SE) vertically below <above> location (i, j).
2. Push the SE up <down> until it touches the underside <topside> of the image surface at some point.
3. Store the pixel value at location (i, j) in S' as the pixel value in S which the SE first touches. This is the minimum <maximum> value, with respect[5] to the SE, in the local neighborhood of the image.

End.

The following two sections, 3.3 and 3.5, discuss the application of morphological processes to the problem of image enhancement.

3.3 The Rotating Rod Algorithm

The objective of this research was to enhance x-rays with the purpose of making nodules more easily discernible. It has been proposed that this objective could only be met with the use of nonlinear techniques. Unfortunately there is no formal recipe for designing an optimal nonlinear process for a specified application. In fact, the majority of our scientific and engineering education is based on linear techniques, because of their relative simplicity and general applicability.

Morphological image processing is a nonlinear approach to image processing and, as such, lacks formal methods for determining the most suitable *filter* for a particular application. According to Song and Delp [26], "A systematic scheme for designing a complete morphological filter using multiple structuring elements is not yet available. Current design is usually based on trial and error." This is in contrast to linear techniques where well documented recipes for the design of linear filters[6] are available.

In the search for designing an appropriate morphological process for the enhancement of

[5]In the case of a three dimensional structuring element (SE), for example: the sphere of the rolling ball algorithm, this minimum <maximum> value is not *necessarily* the minimum <maximum> value of the image in the local neighborhood. Rather, it is the first point on the image which the surface of the SE touches as it is pushed towards the thin undulating image surface. Clearly, this point on the image *will be* the minimum <maximum> when the SE is two dimensional, for example: a disc.

[6]Examples of linear filters for which there are design methods include bandpass filters, notch filters and matched filters.

nodules, the rotating rod algorithm was conceptualized.

3.3.1 Why the rotating rod?

Consider the signal A in Figure 3.3. It is desired that the two bumps (nodules) be extracted from the signal.

Consider moving a window through the signal. The window is chosen to be wider than the bumps. Let the window be five pixels in width (the bumps in the diagram are three pixels in width). The origin of the window is taken as the centre pixel. We now move this window through the signal and compute the minimum value of the signal in the *scope of* (pixels seen by) the window. This minimum value is written to an intermediate signal at the current origin of the window. The intermediate signal would appear as B in Figure 3.3.

If this intermediate signal was now subtracted from the original signal, the output signal would appear as C in Figure 3.3. This is almost the desired output, but the edges of the plateau have been enhanced too. How can this flaw be overcome? These unwanted artifacts will appear regardless of the size of the window; in fact they will be proportional in width to the window dimension.

Consider using the leftmost pixel of the window as the origin. If this window is now moved through the signal a profile, as in D, is obtained. The difference signal, E, is closer to the desired output but the four rightmost pixels of the plateau have been enhanced too. Now, we move a second window through the original profile, this time with its origin as the rightmost pixel. This produces the profile in F.

Profiles F and E are very similar, although they each have failed to entirely reconstruct the plateau. A combination of F and E, however, can regenerate the plateau completely. A new signal, G, is produced as the maximum value per pixel from each of D and F. This new signal G is then subtracted from A yielding exactly the desired output, H.

To summarize the algorithm in one dimension (1D); At each pixel position two minima are found. These minima result from the windows with leftmost and rightmost origins (represented as profiles D and F). The pixel value in the intermediate signal is taken as the maximum of these two minima. This produces the intermediate profile G, which is subtracted from the original A to yield the enhanced profile H.

The above illustrates the operation of the conventional rotating rod in 1D. The *rotating* in rotating rod is there because of the operation of this algorithm in two dimensions. In two

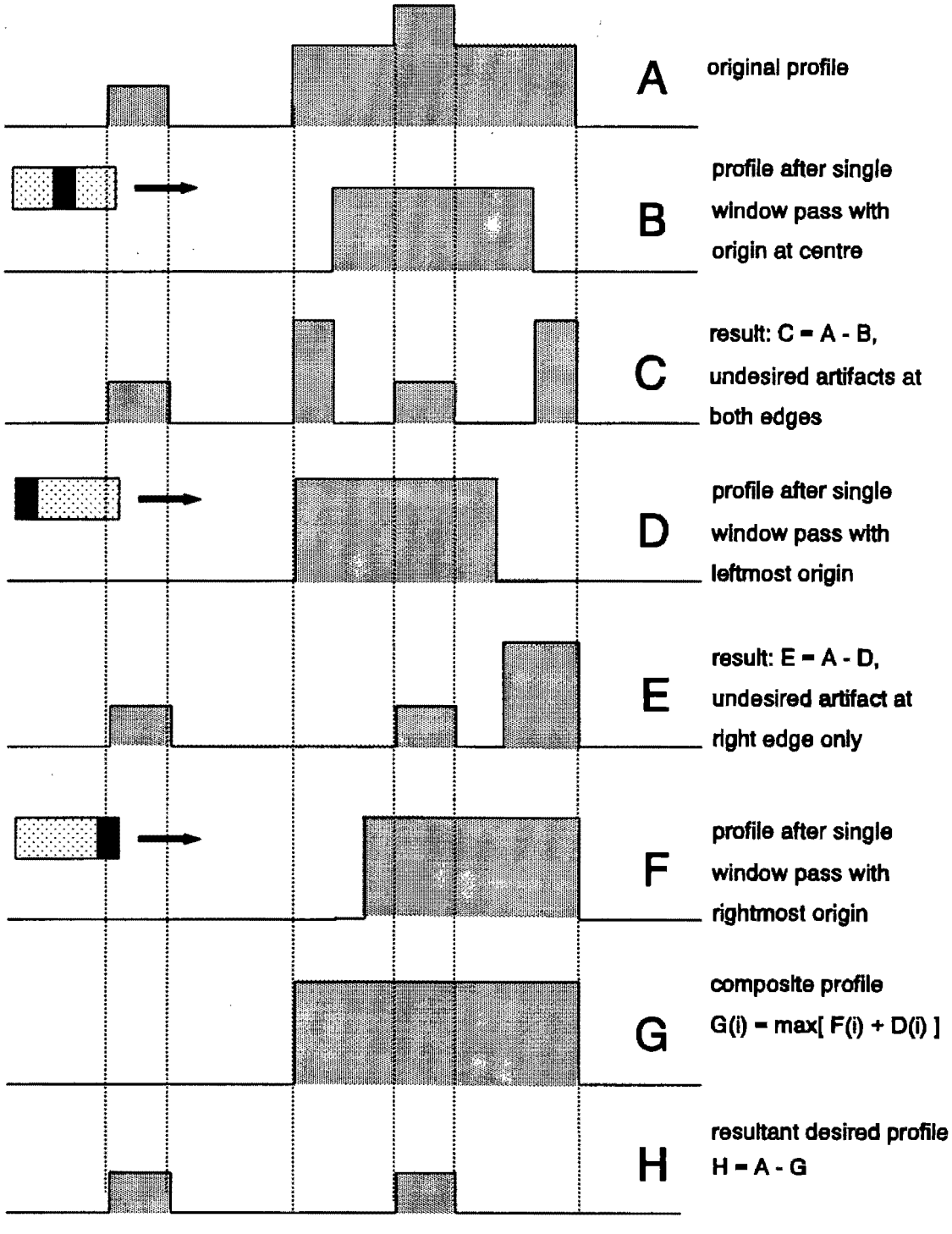


Figure 3.3: Demonstration of the operation of the rotating rod algorithm in one dimension

dimensions the same principle is applied, but the windows are rotated through a sequence of angles. At each orientation two minimum values are found, as in the 1D case (see Figure 3.4). Thus for each pixel in the image S a set of $2n$ minima[7] are found. The maximum value in this set is taken as the new value for that pixel in S'. This generates the intermediate image S'. The output image Y is then produced as the difference image, original minus intermediate.

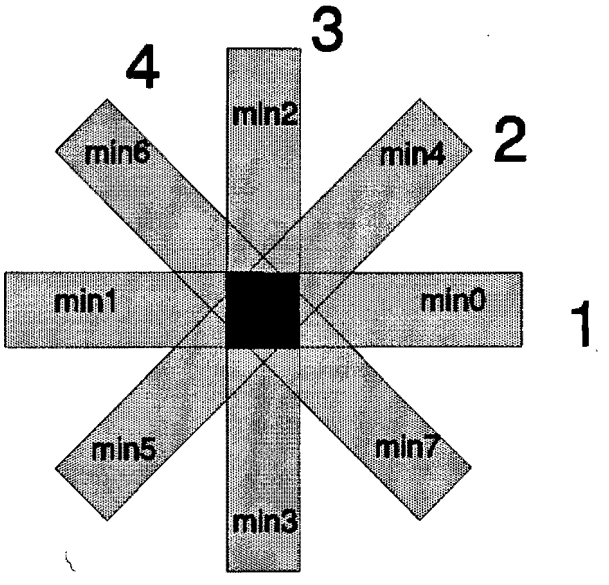


Figure 3.4: The architecture of the 2D rotating rod.

In 1D the rotating rod has the effect of preserving only those structures which are smaller in width than the length of its windows. In 2D the rotating rod has an analogous effect. In the 2D case, however, both the vertical and horizontal dimensions of a structure must be less than the window size. In fact, if the rotating rod operates in more orientations than horizontally and vertically, $n > 2$, then the size of the structure in all of these orientations must be less than the window size, if it is to be enhanced.

It should be clear that small circular structures are enhanced at the expense of any elongated structures, for a given window size. This was one of the driving forces in the development of the rotating rod. It has the ability to remove elongated structured noise, like ribs, while preserving small circularly shaped objects, like nodules. The rotating rod eliminates large scale structures without producing any artifacts. If just one of the orientations of the rod is

[7]The n in $2n$ is the number of orientations at which the windows are positioned. A high number of n is desired for maximum suppression of noncircular structures. A tradeoff must be made between the choice of n and the associated computational burden.

able to fit into a structure, then that structure will exist in the intermediate image and will consequently be eliminated in the difference image.

A further point to note is the inherent robustness of the rotating rod. By finding the maxima of a set of minima, noise spikes will be ignored. This is a similar characteristic to that of the median filter. But, like median unsharp masking, these noise spikes will appear in the difference image.

3.3.2 Algorithmic description of the conventional rotating rod algorithm

The operation of the rotating rod is described by the pseudocode algorithm below. The algorithm requires a single parameter R, which is set to be the maximum expected nodule diameter.

For every pixel in the image S do

1. Compute the minimum value on the horizontal line R pixels to the left, of the current pixel[8,9]. Call it min_0.
2. Compute the minimum value on the horizontal line R pixels to the right of the current pixel. Call it min_1.
3. Compute the minimum value on the vertical line R pixels above the current pixel. Call it min_2.
4. Compute the minimum value on the vertical line R pixels below the current pixel. Call it min_3.
5. Compute the minimum value on the diagonal line north-east from the current pixel, and within a radius of R pixels. Call it min_4.
6. Continue in this manner for the other orientations[10].
7. Find the maximum in the set $\{min_0, min_1, min_2, ...\}$.
8. Write this maximum value to the intermediate image S'.

End for loop.
Generate resultant image as $Y = S - S'$.
End.

3.3.3 A limitation of the rotating rod

A major limitation of the rotating rod algorithm arose from an oversight in the development of the algorithm.

The rotating rod was designed with the assumption that the images of nodules would be superimposed on approximately constant intensity backgrounds. In other words, a best fit line at any of the orientations, in the region of the nodule, would be constant in the intensity dimension. This is a serious oversimplification. It is more often the case that the background is not of a constant intensity (noise assumed negligible).

If the slope of the background is close to zero then the profile through a nodule exhibits a positive slope followed by a negative leeward slope. The nodule is thus still *visible* to the rotating rod. As the slope of the background becomes steeper, the leeward slope of the nodule become less. Above a certain background gradient the leeward slope of the nodule becomes positive, which has the effect of *hiding* the nodule. This is illustrated by Figure 3.5.

Why is the nodule *hidden* from the rotating rod?

The rotating rod has the effect of fitting into sloping surfaces completely. This is understood by investigating the one dimensional case. Consider an image profile with a positive slope (increasing in intensity from bottom left to top right) as in the steep slope profile of Figure 3.5. The two orientations in this plane are west and east. The minimum value for the east orientation[11] is the value of the leftmost data point which is the current pixel. This minimum value is also the maximum of the set of the west and east minima. Thus the current pixel intensity in the original profile is exactly replicated in the intermediate profile. This results in sloping profiles being completely regenerated in the intermediate profile. Consequently the output profile is zero in the region of sloping profiles.

---

[8]The current pixel is also included in the set from which the minimum is determined. This applies for all orientations.

[9]This determination of the minimum value in the scope of the structuring element (SE) may be conceptualised as pushing the SE up underneath S until the SE makes contact with the underneath of S at some point. Also, because the SE for the rotating rod is a line which is a constant in the Z dimension, that point of contact with S gives the minimum value directly.

[10]Orientations { N, S, W, E, NE, NW, SE, SW } are most often used. West, or W, represents the orientation described as horizontal to the left in the pseudocode. An implementation using the additional orientations of { NNE (north north-east), ENE, NNW, WNW, SSE, ESE, SSW, WSW } has also been used.

[11]Recall that the scope of the east orientation is R pixels to the right of, and including, the current pixel.

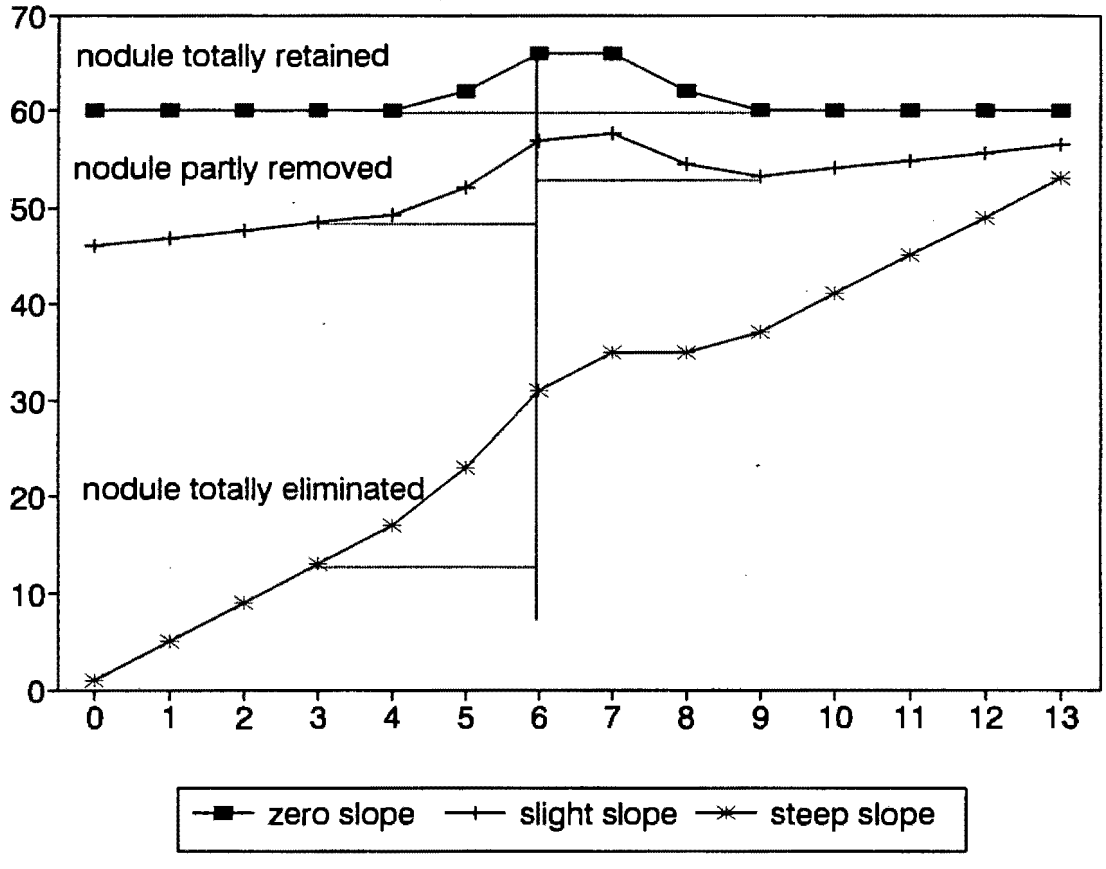


Figure 3.5: A steep background slope has the effect of *hiding* the nodule. The horizontal dotted lines represent two of the orientations of the rotating rod. The degree of removal of the nodule is indicated by the dotted line on the right moving closer to the original profile. Recall that the dotted lines move up and are constrained when they intersect the profile.

By extension, it can be shown that sloping surfaces are also regenerated in the intermediate image, and consequently eliminated in the output image. This is due to the orientation parallel[12] to the gradient giving the maximum value, of the set of minimum values, and that maximum value existing at the current pixel.

Now, reconsider the scenario of a nodule image superimposed on a sloping background. When the leeward slope is positive, the nodule profile is that of a changing positive slope. The rotating rod fits into a positive slope completely, creating a background image containing the nodule. Subtraction to yield the output image results in the elimination of the nodule.

[12]Orientations exist only in the XY plane. If the gradient of the slope is found as a vector of (x, y, z), z being the dimension of intensity, then the parallel orientation is that orientation which is closest, in direction, to the resultant of the x and y components of the gradient vector. Although orientations come in pairs (for example, E and W, or NW and SE), there is only one orientation parallel to the gradient. The other member of the pair is antiparallel.

3.3.4 Modifications to the rotating rod

The limitation, discussed above, may be tackled with two possible approaches.

(a) The first approach is to remove the underlying sloping surface from the image in a two dimensional sense. Such a strategy would then be used as a preprocessing stage to the rotating rod.

(b) The second approach is to compensate for the sloping surface at each orientation of the rotating rod, i.e: in a one dimensional sense.

The solution to both approaches, (a) and (b) is a process which is capable of generating an intermediate surface (as in (a)), or an intermediate profile (as in (b)), which approximates the underlying background information. This is, however, the very crux of the enhancement problem. How does one distinguish background image information from the required image detail? This problem was also addressed in the introduction of section 2.3. Approaches using (a) and (b) are discussed in the following two subsections.

3.3.5 Preprocessing for the rotating rod

Any enhancement module which results in the underlying sloping surface being removed is a potential preprocessor to the rotating rod. In other words, if the enhancement technique satisfies the requirements of additive invariance and structured noise removal then it is a candidate as a preprocessor. From section 2.4 we have seen that unsharp masking (both linear and nonlinear) satisfies these requirements. Later, in section 3.5, we will see that a greyscale top-hat transform is another potential candidate as a preprocessor to the rotating rod.

3.3.6 One dimensional slope compensation for the rotating rod

The second approach (b), above, involves generating a new curve, for each orientation, which best describes the profile of the background surface on which the small scale structures are superimposed. This new curve can then be used as a reference level. The new curve is now raised, conceptually, beneath the original profile until it touches at some point. The analogy of the minimum point in the conventional[13] rotating rod is the value at the origin (location

[13]Conventional refers to the implementation of the rotating rod as described in the algorithm of section 3.3.2, i.e: without preprocessing or one dimensional slope compensation.

of the current pixel) of the shifted data set. This value will always be less than or equal to the original value at the current pixel. The principle of slope compensation is illustrated with Figures 3.6 and 3.7. Figure 3.6 demonstrates the shortcoming of the conventional rotating rod when a nodule is superimposed on a sloping surface, while Figure 3.7 shows the result that is possible with ideal one dimensional slope compensation.

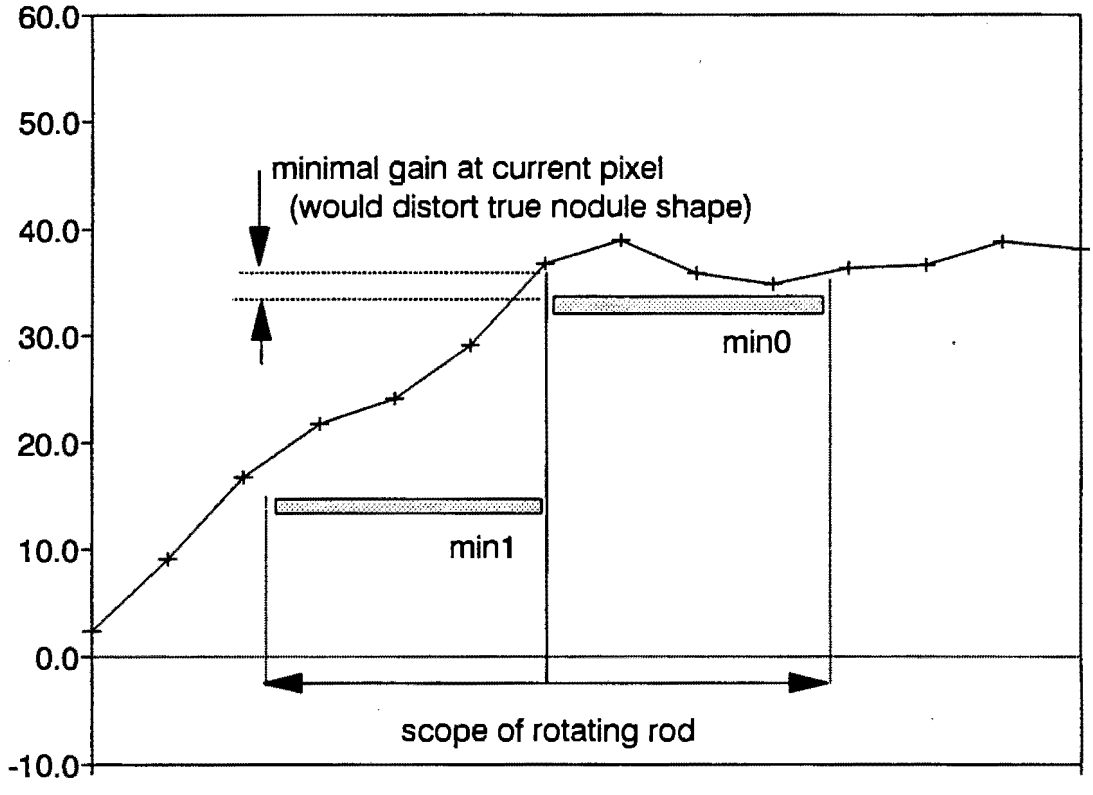


Figure 3.6: The operation of the conventional rotating rod algorithm fails when a nodule is superimposed on a sloping surface. As illustrated in the diagram, the east orientation (which generates $\min_0$) of the rod fits up into the nodule eliminating it to a large extent.

Various methods were attempted for the generation of the intermediate profile.

Least squares curve fitting

A first attempt at generating the intermediate profile involved using straight line least squares (LS) fits in each orientation (of the rotating rod). This allowed an approximation to the gradient of the surface to be incorporated in the rotating rod algorithm. LS fits were inadequate for two reasons: firstly, the LS fit was influenced by the nodule data; and secondly, a straight

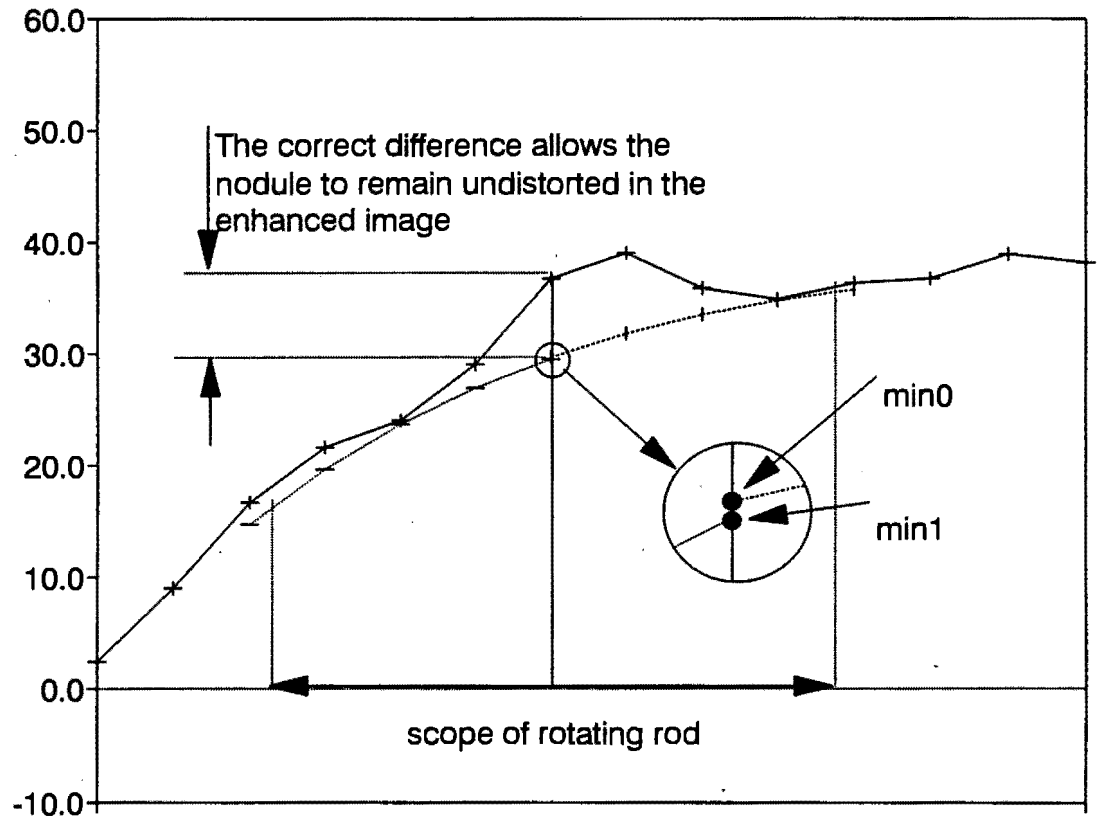


Figure 3.7: One dimensional slope compensation for the rotating rod. min_0 and min_1 are the two orientations of the rotating rod visible in this plane. Each of the slope compensated curves are raised independently beneath the original profile until they intersect the original profile at some point. min_0 and min_1 are then taken as the value of their respective shifted curves at the location of the current pixel. In this diagram min_0 would be used as the replacement value for the current pixel because it is marginally greater.

line was incapable of tracing the background profile acceptably. This is particularly true at regions of discontinuity, for example: step changes.

Two reasons which make the straight line LS fit inappropriate were identified above. The obvious answer to the second problem would be to use a high enough order polynomial LS fit so as to ensure that step changes were suitably approximated. Unfortunately a high order polynomial also allows points of curvature which could trace small scale structures.

Another disadvantage of least squares fitting is the degree of enhancement of small scale structures on the profile is dependent on the degree to which the profile approximates the type of curve (straight line or quadratic, etc.) being fitted. This is a contravention of the *additive invariance* criterion in section 2.2.

Filtering - linear and nonlinear low pass

A simple averaging process may be used to produce a curve which estimates the background profile. However, averaging (linear low pass filtering) is unable to preserve the sharp corners of profiles containing step, or near step, changes. The profile formed by averaging is thus inadequate because the smoothed corners prevent this profile from being raised to the desired level underneath the original data. Consequently the resulting value, taken as the origin of the shifted curve, is inaccurate.

Nonlinear low pass filtering, specifically median filtering, is appealing because the filtered data can accurately reconstruct step changes. Equally attractive is the ability of the median filter to generate the underlying profile without small scale structures and without contributions from those small scale structures. This is subject to a suitable choice of window size for the median filter[14].

Median filtering has the disadvantage, in this application, of being polarity invariant. In other words, small negative, as well as positive, deviations are excluded in the intermediate profile which is generated. In the case of small positive deviations, this is exactly what is required. The intermediate profile will be raised under the original profile until it makes contact at all points except where the deviation is. Narrow negative deviations (valley-like structures in the image) are entirely eliminated when using median filtering as slope compensation. The elimination of negative structures is illustrated by Figure 3.8. This is not necessarily a problem because nodule structures are always positive structures. It does, however, produce unsightly black regions in the output image.

Erosion and dilation for slope compensation

Standard morphological processes of dilation and erosion were implemented in an attempt to generate the underlying profile.

The basic idea is to erode the profile to the stage where the small scale structures have been eliminated, and then to dilate the profile to its original size without the small scale structures. Such an erosion followed by a dilation constitutes a opening, as defined in section 3.2.3. This process is conceivable because of the irreversibility of these morphological processes.

Note that it is a one dimensional data set, the image profile, which is to be opened. The

[14]The window size of the median filter must be greater than twice the width of the structure.

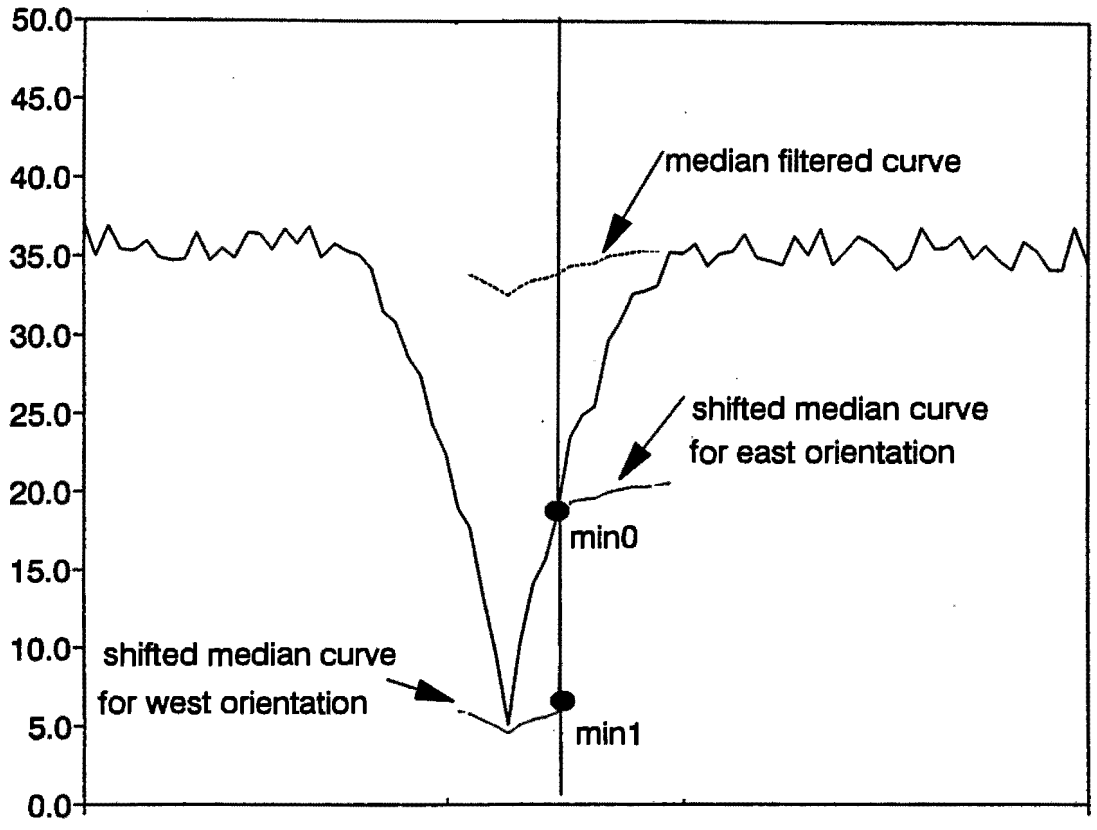


Figure 3.8: The elimination of narrow negative structures when using median filtering as the means of slope compensation for the rotating rod. min_0 is the maximum of the two minima and will be used in the intermediate profile. min_0 is identical to the current pixel value and therefore this current pixel will be set to zero in the output (difference) image.

opening can thus be implemented as one dimensional greyscale opening[15] or as a two dimensional binary opening. The binary image is represented in Euclidean 2-space where a point $p(x, y)$ is set **on** if, and only if, $y \leq f(x)$; where $f(x)$ is the greyscale profile. A binary image generated in this way is the one dimensional analogy of the umbra.

Slope compensation using morphological operators will be discussed in terms of the binary image. Binary erosion replaces an **on** pixel by an **off** pixel if any one of its eight neighboring pixels are **off**. Conversely, dilation replaces an **off** pixel by an **on** pixel if any one of its neighboring pixels is **on**. These operations assume a square 3x3 structuring element.

[15] A one dimensional greyscale opening differs from the standard (two dimensional) greyscale opening in the concept of a neighbourhood. In one dimension the neighbourhood is a number of pixels on either side of the current pixel.

Erosion of the binary image was iterated[16] until the small scale structures were eliminated. The number of iterations is directly proportional to the width of the structures[17]. See Figure 3.9A for an illustration of the binary erosion and dilation processes. An equal number of dilation iterations as erosion iterations are used.

The binary opening was successful in removing small scale structures superimposed on a constant background. Unfortunately, the opening was incapable of effectively eliminating small scale structures from sloping surfaces. Binary erosion was unable to completely remove nodule data from the profile. After a certain number of iterations a non-reducible profile was attained. Any further erosions would not alter this profile. If the nodule data is not entirely removed, the little that is left is regrown in the dilation process. This is illustrated in Figure 3.9B where a nodule, superimposed on a sloping surface, is eroded and dilated.

Non-reducible bumps in 1D slope compensation could be solved by using an area (as opposed to a line) about each rod orientation and then opening that area. This, however, turns out to be the same as performing a standard greyscale opening on the entire image and then using the rotating rod.

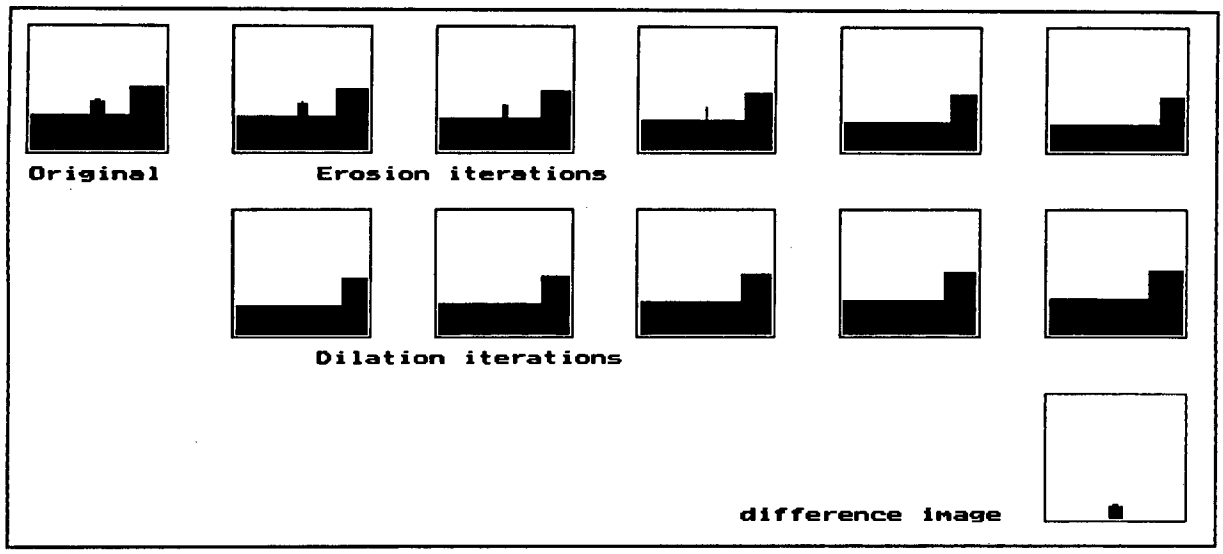
3.4 From the Rotating Rod to Greyscale Opening

One dimensional slope compensation for the rotating rod using binary morphological techniques was investigated. The use of binary morphological operators is easy to understand, and to visualise (see Figure 3.2).

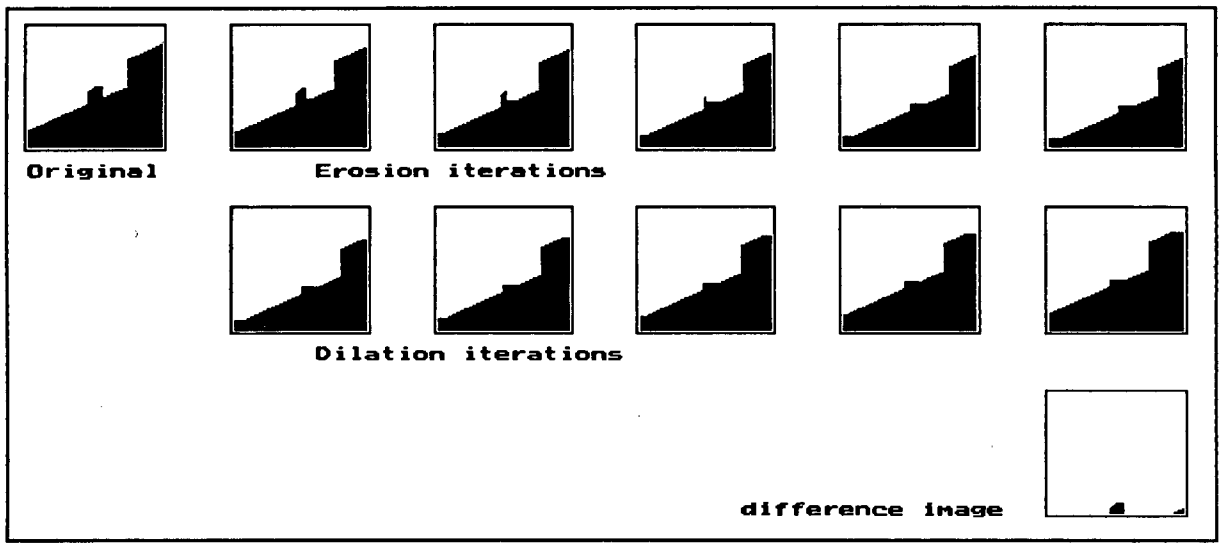
Binary erosion may be thought of as the *peeling* off of a thin layer, from the image, with each application. Binary dilation, on the other hand, applies a thin layer to the image with each iteration. Thus an erosion would remove fine structure, leaving the large predominant structures. An interesting notion is a Mandelbrot figure, where the effect of iterative binary erosion, on the pattern, would be to progressively *eat away* at the fine scale structure revealing the large scale features. It seems logical to extend this idea of successive detail removal into three dimensions. For a crude example, imagine one of those World War II underwater mines - shaped as spheres with protruding detonators. A number of erosion iterations would remove the detonators, leaving the sphere. The process of erosion would also decrease the radius of the mine because of the *peeling* action. The next step is to dilate the mine and effectively

[16]Increasing the size of the local region is equivalent to increasing the number of iterations. This will be discussed further in section 3.5.

[17]The number of erosion iterations should be just greater than than the radius of the nodule to ensure it is totally eliminated.



A.



B.

Figure 3.9: **A**. A sequence of binary erosion iterations on the original image gradually eliminates the nodule. This intermediate binary image is then dilated to raise the background surface to its original height. The difference image contains the undistorted nodule. **B**. This example shows the identical nodule (as in **A**) superimposed on a sloping surface. Erosion iterations are unable to entirely remove the nodule because a non-reducible shape is formed. Consequenlty the difference image shows a distorted nodule; it should have appeared as in **A**.

regrow it to its original radius. Now if we were to subtract, conceptually, the eroded mine (sphere only) from the original mine, we would be left only with the detailed information (detonators).

After some thought about this idea of eroding in three dimensions it should become clear that greyscale erosion does almost exactly that. Imagine that the mine, discussed above, was quantised into voxels. If we were to replace each voxel by the minimum intensity[18] voxel of its immediate neighbours we would erode away at the detonator from all sides. The difference between finding the minimum value and *peeling off*[19] is that constant surfaces do not shed a layer when the minimum neighbourhood value is used to replace each voxel. This difference is insignificant in a morphological opening because the regrowth, or dilation, process is entirely complementary. In other words, the opened surface will be identical regardless if it was generated by min max methods or by peeling and reapplying. Of course, the intermediate image (or mine) will be slightly different.

The informal argument above makes it seems likely that a greyscale opening would be appropriate for generating an image which has been stripped of its detail. It remains to subtract this *opened* image from the original in order to produce an image containing the detail only.

3.5 Greyscale Opening and the Top-hat Transform

A greyscale opening is a composite operation of a greyscale erosion and a greyscale dilation.

As described in section 3.3 binary erosion is achieved by replacing an **on** pixel by an **off** pixel if any one of its eight neighbouring pixels are **off**. Extending this binary concept to greyscale involves replacing each pixel by the minimum value in its neighbourhood. Several iterations are often required to erode small scale structures sufficiently. However, increasing the size of the local region, from which the minimum is found, is identical to increasing the number of iterations[20].

[18] Intensity in this example would be measured as radial distance from the centre of the mine.

[19] This argument is somewhat informal. Careful consideration of the actual implementation of *peeling off* in two, or three, dimensions reveals the discrete nature of the problem. Simply turning an **on** pixel **off**, if any of its neighbours are **off**, is not an accurate implementation of our conceptual understanding of *peeling off*, especially if this process is iterated. *Peeling off* is better modelled by inspecting a radial neighbourhood. Extending this to three dimensions implies that a spherical region is the most suitable neighbourhood if *peeling off* is to be simulated. This is the well known *rolling ball* structuring element (to be discussed in section 3.5).

[20] The relationship between number of iterations and footprint dimension has, as yet, only been validated for the simple flat-topped square structuring element.

This description of greyscale erosion made no mention of a structuring element (SE). The effective SE, for the process described above, was a flat-topped square. Using more complex SEs involves inspecting the distances between the image and the SE for each image pixel in the scope of the SE. The current pixel is then replaced with the intensity of the image pixel which is *closest* to the SE.

Greyscale dilation is the complement of greyscale erosion. The eroded image is subjected to greyscale dilation to produce the opened image. Opening X by structuring element B is denoted as $X \circ B = (X \ominus B) \oplus B$. The enhanced image is generated as the difference image, original minus opened image. Thus if the original image is X, then the enhanced image is given as:

$$X_{enhanced} = X - (X \circ B)$$

An image enhanced by this process, original minus opened, is said to have been *top-hat transformed* [18, pp701].

3.5.1 Simple greyscale opening

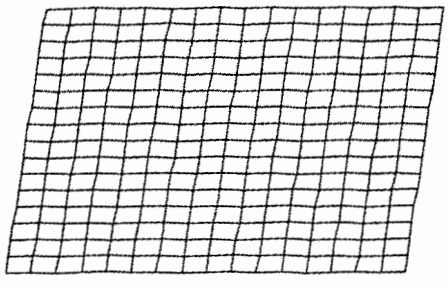
The simplest opening is that which uses a flat-topped square as a SE. The operation of an opening using such a SE degenerates to generating an intermediate image by finding the minima in each local region. This intermediate image is then processed to produce the opened image by finding the maxima in each local region.

Openings using a disc, a hemisphere, a parabaloid and a cone as the SE have also been investigated. These structuring elements are shown in Figure 3.10.

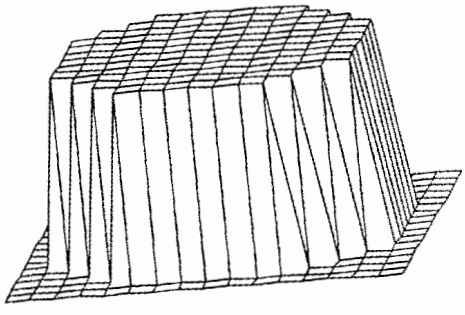
3.5.2 The rolling ball algorithm

Greyscale opening, or closing, is called the *rolling ball algorithm* when the structuring element is a sphere[21]. This technique has proved successful in applications ranging from angiogram enhancement [1] to the segmentation of the surface froth structures in flotation cells [31]. Based on the discussion of section 3.4, iterative applications of the rolling ball algorithm on an image may be termed *peeling* the image.

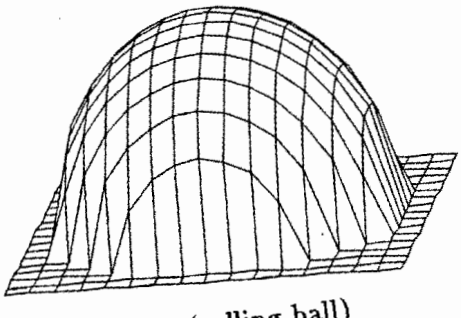
[21] It is sufficient to use a mathematical model of a hemisphere to simulate a sphere.



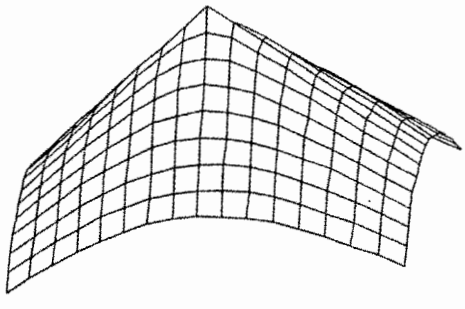
A. Flat Square



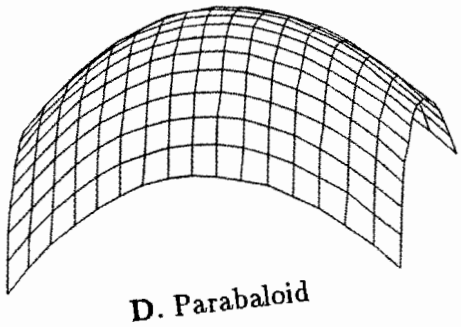
B. Flat Disc



C. Hemisphere (rolling ball)



D. Cone



D. Parabaloid

Figure 3.10: Illustration of common structuring elements

3.5.3 Examination of the greyscale openings

The top-hat transform using the rolling ball is intended as a detail enhancement process and/or a dynamic range reduction process. With regard to detail enhancement the rolling ball may be inappropriate if it *fits* into nodule structures in the image. This is because the nodules themselves approximate a hemispherical mound in the image.

This fitting of the rolling ball SE into the nodules is undesirable because S' contains these structures and consequently Y does not. The solution is to use a ball large enough so as not to fit into nodule structures. A large ball has the disadvantage that intermediate structures, i.e: smaller than the ball but larger than the nodules, are not eliminated. Remember that increasing the size of the ball decreases the curvature of the ball which prevents the ball from fitting into the nodule structures. A flat-topped square, or disc, has zero curvature regardless of its dimension. In other words a square, or disc, just larger than the nodules appears to be more suitable than a ball because the ball would have to grow unacceptably large in order for its curvature to be sufficiently diminished.

The size of the SE has an important bearing on the resultant image. With a flat-topped square all objects of lesser dimension, in any orientation[22], will be enhanced. Thus increasing the size of the SE increases the sizes of objects which will be enhanced. For maximum discrimination the SE is designed to be just larger than the size of the sought after structures. In other words, if a nodule with a diameter of five pixels is to be enhanced, a SE of 7x7 is appropriate. The dimension of the SE is always an odd number of pixels to ensure symmetry.

Operation of a simple top-hat transform on the test image indicates that the top-hat transform is additively invariant (see Figure 2.3).

3.6 A Comparison of the Rotating Rod and the Top-hat Transform

The effective scope of all the rod orientations, used in the rotating rod algorithm, and the disc structuring element (SE), used in the top-hat transform, is the same. However, there is a subtle, but significant difference in the operation of these algorithms.

[22]Contrast this to the rotating rod where an object must be smaller in *all* dimensions for it to be enhanced. This is discussed in more depth in the next section.

Consider "islands" of various shapes and sizes in a binary image[23]. Imagine that the disc SE is partitioned into sectors, where each sector corresponds in orientation to the rods of the rotating rod. Allow each sector to collapse to a single pixel line at each orientation. If this partitioned SE is to behave identically to the original disc SE it must obey the following rule: the binary output value must be 1 if any of the sectors do not fit; it will be 0 only if all the sectors fit. This is stated formally below.

Let F denote a boolean variable which has the value 1 if the sector fits[24] into the image structure, and a value 0 if the sector is restrained at any point from fitting in completely. At every position in the binary image, the binary output value is given as

$$\neg F_1 \vee \neg F_2 \vee \neg F_3 \vee \cdots \vee \neg F_n$$

if there are n orientations. The $\vee$ and $\neg$ represent the logical OR and logical NOT operators, respectively.

By the definition of the rotating rod algorithm, the binary output value, for an image processed by the rotating rod, is

$$\neg F_1 \wedge \neg F_2 \wedge \neg F_3 \wedge \cdots \wedge \neg F_n$$

where $\wedge$ is the logical AND operator.

Thus, the top-hat transform, with a disc as the SE, will retain any island if the size of the island in **any** orientation is less than the radius of the disc. Conversely, the rotating rod algorithm will only retain an island if the size of the island is less in **all** orientations than the length of the rod. In summary, the rotating rod will eliminate elongated islands if the length of the island structure is longer than the rod length, whereas the top-hat transform will retain elongated islands if the width of the islands is less than the radius of the SE disc.

3.7 Summary of the Morphological Enhancement Modules

The morphological enhancement modules are summarized in terms of the requirements described in section 2.2. A table, 3.1, for the evaluation of additive invariance, structured noise elimination, artifact production, rotational invariance, robust operation and processing time is provided below. As in section 2.4, processing time is written in order notation, followed by

[23] A binary image is considered, as opposed to a greyscale image, to avoid unnecessary complexity. The concepts discussed in this section are equally applicable for greyscale images.

[24] For a sector, or a rod, to fit into an image structure, it implies that the binary image has the value 1 for all pixels in the scope of that sector, or rod.

a label which represents the predominant type of operation. The interpretation of processing time is the same as in section 2.4.

Enhancement modules	Additive invariance	Removal of structured noise	Artifact production	Rotational invariance	Robust operation	Processing time
Rotating Rod Algorithm	NO	YES - uses background subtraction	NO[25]	YES	fair	$O(n^2)$ boolean
Top-hat Transformation	YES	YES - uses background subtraction	unknown - appears not	YES	NO	$O(n^2)$ integer

Table 3.1: Comparison of morphological enhancement techniques.

With regard to enhancement for best visual representation, reduced dynamic range and detail enhancement are also taken into consideration. The morphological techniques use background subtraction and thus provide dynamic range reduction, for the most part of the image. Isolated regions of high intensity might still exceed the required dynamic range.

Although the simple top-hat transform preserves only those structures which are smaller than the footprint of the SE, it is possible, in principle, to do better. With a top-hat transform, all anatomical detail which is smaller than the SE will be enhanced. All of the enhancement modules discussed thus far suffer from the problem of enhancing anatomical detail.

There exists another technique which has not been extensively investigated. This technique is called image warping. It operates by warping a reference image[26] to register with the image to be diagnosed. The warped reference image is then subtracted from the x-ray under inspection. The difference image, thus generated, should, in principle, contain only information which is indicative of some pathology. Naturally *normal* anatomic detail (like blood vessels) will appear in both images and should therefore not be enhanced.

[25]The rotating rod does not generate artifacts, but it does remove information.

[26]This image is chosen to be representative of the average chest x-ray and to be clear of any pathology.

Chapter 4

Image Analysis

4.1 Introduction

The image analysis module in the automated nodule detection system is concerned with processing an input image to produce data which may be classified. A common form of image analysis is segmentation. An image can be segmented into two or more regions based on certain characteristics. In this application we are interested in detecting the locations of possible nodule sites in an image. The nodule in the image can be reasonably accurately modelled. There are, however, some factors which result in a loss of model accuracy. These factors will be discussed in section 4.4.6 where the analytic nodule model is developed.

The ability to model the objects being detected and the uncertainty of the locations of these objects make this problem suitable for the technique of template matching. Template matching, using a model of the object to be detected, is the *brute force* approach according to Rosenfeld [10]. Rosenfeld goes on to introduce the Hough transform as an alternative approach to finding specific patterns in an image.

In this chapter the requirements of the image analysis module are discussed. Both the Hough transform and a variation of template matching require an edge detected image. Before investigating template matching and the Hough transform, edge detection as a preprocessor to the image analysis module is discussed. Then the main sections of template matching and the Hough transform are presented. Finally a summary of all the image analysis modules is made.

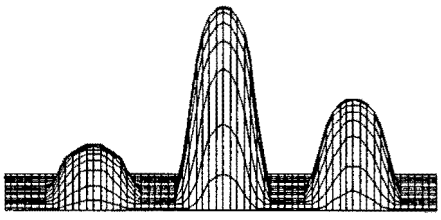
4.2 Requirements of the Analysis Module

Template matching is used as an image detection procedure which produces a surface with peaks corresponding to the centres of suspected objects. The Hough transform, an alternative image analysis procedure, also produces a surface (or surfaces) where peaks correspond to the centres of suspected nodule sites. For the application of nodule detection in chest x-ray images the image analysis module should satisfy the following requirements:

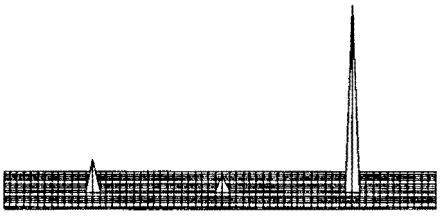
- **Additive invariance**: the processing of identical objects in an image should produce an equal output, in shape and in intensity, regardless of the position of the object in the image, and regardless of the average energy in that local region of the image.
- **Robust operation**: the analysis operation should be insensitive to noise.
- **Short processing time**: for the same reason as that mentioned in section 2.2, processing time should be minimised.
- **Good discrimination**: the output for a hit should be easily distinguishable from a non-hit.
- **Invariant measure of match**: allows comparison of the *goodness* of a match regardless of the template[1] and the image.
- **Intensity matching**: for this application the intensity (vertical) scale of the object being detected is as important as the shape description and should therefore be included in the measure of match. This requirement differs from the standard correlation techniques, where output intensity is either proportional to input intensity or invariant to input intensity.

  An example of intensity matching is provided in Figure 4.1. Figure 4.1A shows three nodule models which are identically shaped but differently scaled in the intensity dimension. If it is desired that objects of the intensity of the right hand side (RHS) model are to be detected, then we would hope for an output, after template matching, which resembled 4.1B. The template matching section, 4.4, will show that many correlation-type template matching procedures actually produce an output as in Figure 4.1C. In other words, these procedures are intensity invariant.

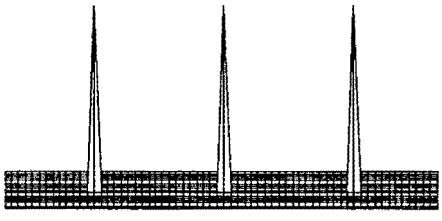
[1] An analogous concept to the template exists when using the Hough transform. The template may be likened to the algorithm which determines the filling of the Hough accumulator array.



A. Test image containing 3 nodule models



B. Resultant surface after using intensity matching for the RHS nodule



C. Resultant surface without using intensity matching

Figure 4.1: Illustration of intensity matching in template matching

4.3 Edge Detectors for Image Analysis

Three edge detectors were investigated for use in the image analysis module. The motivation for the use of the chosen edge detectors is understood by the requirements of the edge detection routine in this context. In general, edge detectors are very sensitive to noise because of their inherent differentiation. The sensitivity of the edge detectors does not only amplify noise, but often it also results in true edges being poorly defined in the edge detected image. For the application of preprocessing to subsequent image analysis it was required that the edge detection process be robust and produce crisp edges.

These requirements are similar to Canny's [4] edge detection performance criteria of:

- Good detection. By maximising the SNR there should be a low probability of obtaining false edges and a low probability of falsely enhancing a nonedge.
- Good localization. The detected edge should be as close as possible to the centre of the true edge.
- A single response for a single edge. Only one dominant peak should be present, for a single edge, in the edge detected output.

The edge detector which was first attempted was the well known Sobel operator. The Sobel operator computes a measure of the gradient in the horizontal and vertical directions. The output intensity at each position is then given as the vector sum of the horizontal and vertical gradient strength. This edge detector was found to be too sensitive to noise, with the result that undesired, or false, edges were enhanced.

The second edge detector was the difference of areas edge detector (DOAED). The DOAED was investigated because of its tunability, its noise robustness, and its history of success in this application [3]. The DOAED uses two adjacent windows of a given size. The window size is a user defined parameter, and it this parameter which allows the DOAED to be tuned. The one dimensional summations of the intensity values of the pixels within each window are computed. The output at the centre pixel value is the difference of the two summations. These windows are oriented vertically and horizontally, and centred on the current pixel. The vector magnitude of the horizontal and vertical outputs is taken as the edge detected value for that current pixel.

This edge detection process can be mathematically described as the convolution of the edge detector windows with a data stream, *viz.* the pixels. The edge detector can be analysed using linear filter theory. The one dimensional description of the filter is:

$$h(n) = \sum_{i=-W}^{-1} \delta(n+i) - \sum_{i=1}^{W} \delta(n+i)$$

where W is the width of each window and $\delta(n)$ is the impulse function.

This filter has the disadvantage that all edges are spread by the convolution process. The amount of spread is directly proportional to the window size of the edge detector. It is desired that the edge detector is robust and, at the same time, produces a narrow (unspread) output for all edges. These requirements appear to be contradictory because both the robustness *and* the spread of the DOAED are proportional to the window size. A possible solution to this problem is provided by the non-spread edge detector discussed below.

The third edge detector was a refinement of the difference of areas edge detector called the nonspread edge detector (NED). The NED operates according to the algorithm below:

For every pixel in the image

1. Compute the absolute value of the difference between the horizontal windows about the current pixel.
2. Compute the absolute values of the differences between the horizontal windows centered behind and ahead of the current pixel.
3. Decide whether the current difference is greater than both the difference behind and the difference ahead, *i.e:* is the current difference a local maximum in the horizontal direction?
4. Perform the same computations in the vertical orientation.
5. If either the vertical or horizontal differences are local maxima then write the output value, for the current pixel, as the vector magnitude of the vertical and horizontal components. If neither differences are maxima then write a zero as the output value.

End.

The NED has the advantages of (a) being as robust as the DOAED, (b) being tunable, and (c) generating a single pixel line for an edge.

The operation of the difference of areas and the nonspread edge detector is illustrated by Figure 4.2 which shows three surface plots. The first surface plot shows the model of a nodule. The second plot indicates the output after the application of the difference of areas edge detector. The third surface plot shows the output of the nodule after edge detection with the nonspread edge detector. The two cross sections (profiles) are taken through the difference of areas image and the nonspread image at the same row. These profiles illustrate the narrow response of the nonspread edge detector.

4.4 Template Matching

4.4.1 Introduction

Template matching has been identified as an image analysis technique in the detection of nodules. Template matching is the two dimensional extension of the one dimensional matched filter. To detect nodules using template matching we generate a model of the x-ray image of a nodule - this is the template.

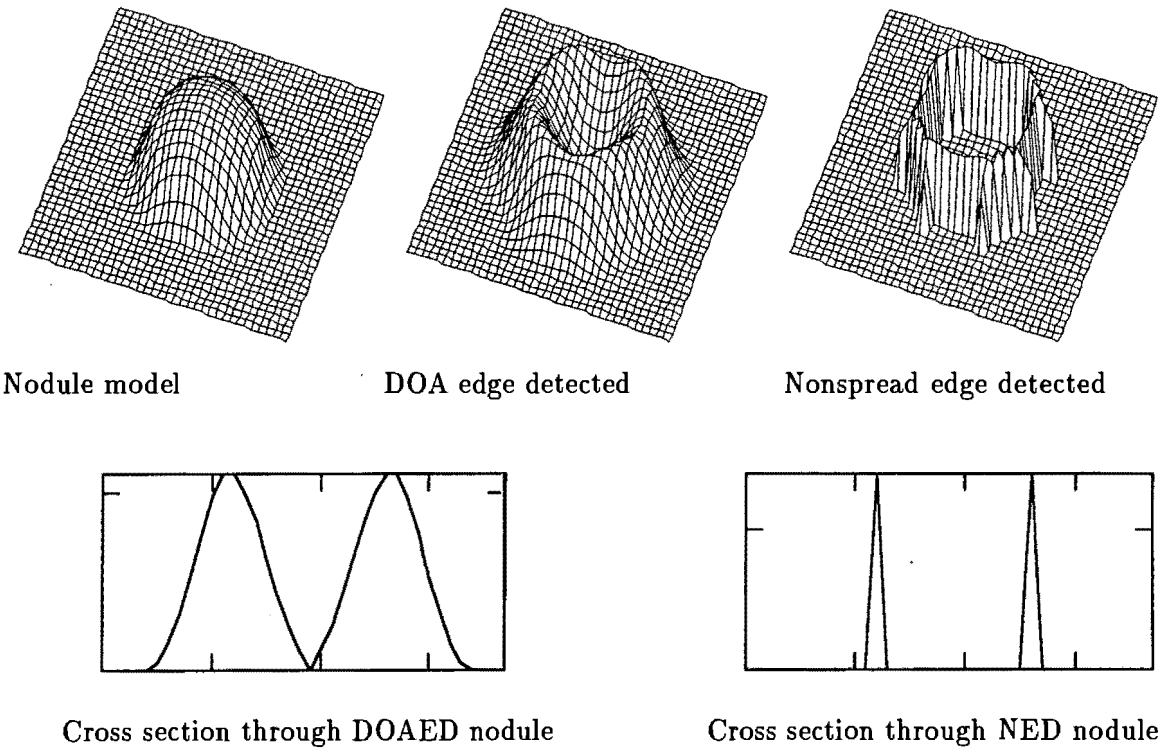


Figure 4.2: Comparison of the difference of areas (doa) edge detector and the nonspread edge detector when applied to a model of a nodule. Cross sections through the edge detected images are also shown.

There are several ways of implementing template matching. Section 4.4.2 discusses the cross-correlation technique and three variations on this theme. An additional implementation of template matching, namely variance matching, is also investigated.

4.4.2 Template matching using a correlation measure

The theory of template matching is an extension of the one dimensional matched filter theorem [29] found in most signal processing texts. The following discussion is that of optimal linear operators as found in [24]. Rosenfeld [24] states at the beginning of chapter 9 that nonlinear operators may give better results. The theory of template matching is developed from first principles by Rosenfeld in [24, chapter 9] and is summarized here for completeness.

A measure of match (or mismatch) between two functions f and g over a local region A may be made using one of several formulae, or a user-defined formula. Commonly used formulae include

$$\max_A |f-g|, \qquad \iint_A |f-g|\,dA \qquad \text{and} \qquad \iint_A (f-g)^2\,dA$$

Using the last formula, we have

$$\iint_A (f-g)^2\,dA = \iint_A f^2\,dA + \iint_A g^2\,dA - 2\iint_A fg\,dA$$

For a given $\iint f^2\,dA$ and $\iint g^2\,dA$, $\iint fg\,dA$ may be used as a measure of match[2]. The same conclusion can be reached using the Cauchy-Schwarz inequality.

In template matching, we assume that the footprint of f, the template, is much less than the footprint of g, and that f is zero outside a small region A. Template matching is then achieved by shifting the template f into all positions relative to g, the image, and computing $\iint fg\,dA$ for each shift.

We have

$$\iint_A fg\,dA = \iint_A f(x,y)g(x+u,y+v)\,dxdy \tag{4.1}$$

where (u,v) is the absolute pixel location in the image.

But, since f is zero outside the region A, the right hand side of the previous equation may be written as

$$\iint_{-\infty}^{\infty} f(x,y)g(x+u,y+v)\,dxdy$$

which is the cross correlation C_{fg} of f and g.

The correlation coefficient r is formally defined [22] as

$$r = \frac{\iint_A fg\,dA}{\sqrt{\iint_A f^2\,dA . \iint_A g^2\,dA}} \tag{4.2}$$

Note that $-1 \leq r \leq 1$, where -1 represents *a perfect negative match* and +1 represents *a perfect match.* A value of 0 is interpreted as zero match between f and g.

[2] $\iint (f-g)^2\,dA$ was used as a measure of *mis*match.

Since $\iint_A f^2\,dA$ is constant, a normalized correlation measure of

$$\frac{C_{fg}}{\sqrt{\iint_A g^2\,dA}}$$

may be used, as is given in [24]. This normalized measure will vary, however, if different templates are used. This is in contrast to the correlation coefficient r of equation 4.2 which is an invariant measure of match.

A larger value of A will, in general, provide a more robust measurement for r because the area of integration has been increased.

4.4.3 Evaluation of correlation template matching methods

The standard technique and three variations of template matching using correlation measures are evaluated in terms of the image analysis requirements. Specifically, the requirements of (i) additive invariance, (ii) match measure invariance and (iii) intensity matching (as discussed in section 4.2) will be investigated for each variation of template matching[3].

Consider matching the template f with the image g. Within g are two potential matches g_1 and g_2. g_1 has the identical mean value as f, but g_2 has a different mean in its local region. g_2 is identical to g_1 in all other respects. We can write g_2 as $g_1 + k$, where k represents the mean difference between g_1 and g_2. We compute the output values at a and b, the centre locations of g_1 and g_2. Additive invariance is then determined by computing the output at location a and the output at location b. An identical output is required for additive invariance to be satisfied.

Simple correlation measure

The simple correlation measure uses equation 4.1.

The template f is designed to accurately model the desired signal. Therefore we can write g_1 as f because they are identical signals within their respective local regions. Thus g_2 becomes $f + k$.

[3]Recall that additive invariance implies an equal output for an object, regardless of its background surface; match measure invariance provides a measure of the degree of match, irrespective of the image and the template; and intensity matching uses the vertical scale information of the object as an additional matching parameter.

At location a the output value is computed as

$$\iint_A f g_1 \, dA = \iint_A f^2 \, dA$$

while at location b the output value is

$$\iint_A f.(g_1 + k) \, dA = \iint_A f.(f + k) \, dA$$

which simplifies to

$$\iint_A f^2 \, dA + k \iint_A f \, dA$$

Uncorrelated high intensity regions, indicated by the value of k, give larger output in the correlation surface than genuine matches. This is not surprising because, according to Duda and Hart [8], "...using cross-correlation in template matching involves a tacit assumption that the picture energy in every window is roughly the same." This assumption is clearly not satisfied, in general.

This method does not give an invariant measure of match. There is also no intensity matching.

Correlation measure with zero-mean template

Correlation measure with zero-mean template uses equation 4.1 with the modification of a zero-mean template, $\iint_A f \, dA = 0$.

At location a the output value is computed as $\iint_A f g_1 \, dA$, while at location b the output value is $\iint_A f.(g_1 + k) \, dA$

The output at b simplifies to

$$\iint_A f g_1 \, dA + k \iint_A f \, dA$$

which reduces to $\iint_A f g_1 \, dA$ because $\iint_A f \, dA = 0$

Thus the output values at a and b are identical and therefore the zero-mean template satisfies the additive invariance criterion. This method does not give an invariant measure of match. Also, there is also no intensity matching; instead output intensity is proportional to input intensity.

Normalized correlation measure

In this section, and in the next, a similar result will be used. For this reason it is developed here in a generic manner. The result sought is the output at location b when the template is unspecified and the normalized cross correlation equation, 4.2, is used. Thus, computing the output value at b we have

$$\frac{\iint_A f.(g_1 + k)\, dA}{\sqrt{\iint_A f^2\, dA . \iint_A (g_1 + k)^2\, dA}}$$

Which expands to the generic form below

$$\frac{\iint_A f g_1\, dA + k \iint_A f\, dA}{\sqrt{\iint_A f^2\, dA . [\iint_A g_1{}^2\, dA + 2k \iint_A g_1\, dA + k^2]}} \tag{4.3}$$

Use of equation 4.1 with normalization gives equation 4.2 which is the normalized cross correlation measure. The template used in normalized correlation measure is designed to match the signal and therefore we can write g_1 as f.

At location a the output value is computed as:

$$r = \frac{\iint_A f^2\, dA}{\sqrt{\iint_A f^2\, dA . \iint_A f^2\, dA}} = 1$$

and thus the correlation coefficient is, as expected, a perfect match.

At location b, using equation 4.3 with $g_1 = f$, we have

$$\frac{\iint_A f^2\, dA + k \iint_A f\, dA}{\sqrt{\iint_A f^2\, dA . [\iint_A f^2\, dA + 2k \iint_A f\, dA + k^2]}}$$

Squaring the numerator, for comparison with the expression under the square root sign, we have

$$\text{numerator}^2 = \left[\iint_A f^2\, dA\right]^2 + 2k \iint_A f^2\, dA . \iint_A f\, dA + \left[k \iint_A f\, dA\right]^2$$

while the denominator squared (dropping the square root sign) evaluates to

$$\text{denominator}^2 = \left[\iint_A f^2\,dA\right]^2 + 2k \iint_A f^2\,dA . \iint_A f\,dA + k^2 \iint_A f^2\,dA$$

So, the quotient is not unity because the last term of the numerator and denominator differ and therefore the normalized correlation measure does not satisfy the additive invariance criterion. This method does give an invariant measure of match, but there is no intensity matching.

Normalized correlation with zero-mean template

Normalization gives an invariant measure of match, while a zero-mean template gives additive invariance. It is hoped that this combination will offer the advantages of both normalized and zero-mean correlation.

Normalized correlation with zero-mean template is described by equation 4.2 with the modification of a zero-mean template, $\iint_A f\,dA = 0$.

At location a the output value is computed as:

$$\frac{\iint_A f g_1\,dA}{\sqrt{\iint_A f^2\,dA . \iint_A g_1{}^2\,dA}}$$

At location b the output evaluates to equation 4.3. But, with $\iint f\,dA = 0$, and distributing $\iint_A f^2\,dA$ in the denominator, we get

$$\frac{\iint_A f g_1\,dA}{\sqrt{\iint_A g_1{}^2\,dA . \iint_A f^2\,dA \;+\; 2k \iint_A g_1\,dA . \iint_A f^2\,dA \;+\; k^2 \iint_A f^2\,dA}}$$

So, the quotient is not the same as at a because of the additional terms $2k \iint_A g_1\,dA . \iint_A f^2 dA$ and $k^2 \iint_A f^2\,dA$ in the denominator. Therefore normalized correlation with a zero-mean template does not satisfy the additive invariance criterion. This method does give an invariant measure of match, but there is no intensity matching.

Summary of correlation-type template matching

The correlation methods investigated as template matching routines are summarised in terms of the requirements of additive invariance, invariance in match measurement and intensity matching (see Table 4.1).

Type of template matching	Additive invariance	Invariant measure of match	Intensity matching
Simple correlation measure	NO	NO	NO
Correlation measure with zero-mean template	YES	NO	NO
Normalized correlation measure	NO	YES	NO
Normalized correlation with zero-mean template	NO	YES	NO

Table 4.1: Comparison of template matching techniques using correlation measures.

From the table it is evident that none of the methods use the intensity of the object being detected as a matching criterion. It is also clear that preprocessing is necessary for all implementations of template matching, except the zero-mean template matching procedure. The zero-mean template method, however, has the disadvantage that it does not provide an invariant measure of match.

4.4.4 Template matching using a variance measure

An alternative scheme for template matching was developed. This method uses the inverse of the variance between the image and the template as a measure of match. It also requires a preprocessed image with a constant background intensity because it does not satisfy additive invariance. This variance matching method has been designed to give an invariant measure of match and to provide intensity matching. For every pixel in the image, we compute:

$$\frac{1}{1+\frac{1}{energy}\iint_A (f-g)^2\,dA} \tag{4.4}$$

where $energy = \iint_A (f-k)^2\,dA$ (k is the intensity level on which the template f is superimposed), giving a value of $\frac{1}{2}$ on a level region of the image. A complete match has the denominator of equation 4.4 reducing to unity, and thus giving a match measure of unity. Alternatively, a match measure of zero arises as the denominator grows very large, as in the case of a large deviation from the template.

4.4.5 Template matching using first differences

Some authors believe that edge detection prior to template matching is a better practical solution.

This is the opinion of Woods [10, pp77-81] where he writes "The matched filter can be derived by maximising a filter's output SNR. A shortcoming of this approach, however, is that it does not obtain the bias correcting energy term. In fact, ... the matched filter thus obtained is more sensitive to signal energy than to signal shape. The advice given there to correct the problem is to use a so called derivative matched filter to emphasize the high-frequency or edge influence on the detection."

Rosenfeld also supports this method. He states: "Because images are usually quite correlated, a copy of the desired pattern is usually not the ideal matched filter, it is better to cross-correlate the first differences of the pattern and the image with each other. So this implies that edge detected template matching is a better proposition" [10, pp268].

A short discussion of edge detectors was provided in section 4.3. The image and template should be subjected to an identical edge detection process. This is to ensure that the template (model of the nodule) is modified in the same manner as the nodules in the image.

Once the image and template have been edge detected, it is still necessary to perform the template matching. We are again faced with the decision of which template matching paradigm to use.

4.4.6 Design of the template

In order to perform template matching, it is necessary to generate an accurate template. The template should best model the object being detected.

The received x-ray energy $I(x, y)$ is a function of both the x-ray attenuation μ of the material through which the beam passes and the thickness of the material at that point $z(x, y)$. Thus, for an incident x-ray energy source I_o, we have:

$$I(x, y) = I_o e^{-\mu z(x,y)} \tag{4.5}$$

The exponential nature of the x-ray equation 4.5 implies that the intensity of a nodule image

will differ as a function of the attenuation[4] of the background tissue. For example, a nodule behind a rib will produce an image of lower intensity than the same nodule obscured by lung tissue only. Processing is largely simplified if the image can be transformed into a domain where attenuation is additive. Such a transform is investigated.

Transformation of the x-ray image into a linear attenuation domain

The measure of x-ray film blackening[5] is the photographic density:

$$D = \log_{10}\left(\frac{L_o}{L}\right) \tag{4.6}$$

where L_o and L are the incident and transmitted light intensities, respectively for the blackened developed film [15]. Density is generally plotted as a function of logarithmic exposure to light, to produce a characteristic S curve. In diagnostic radiology this relation is applicable because x-radiation is converted to light via intensifying screens [15]. In practice the linear portion of the density versus log exposure curve is used [15, pp415], making density proportional to log exposure. Exposure is synonymous with the received x-radiation $I(x,y)$, and therefore we can write

$$D = \gamma \log_{10} I(x,y) \tag{4.7}$$

where γ, which is the slope of the density versus log exposure curve, is a property of the film. Substituting equations 4.5 and 4.6 into equation 4.7 yields

$$\log_{10}\left(\frac{L_o}{L}\right) = \gamma \log_{10}\left(I_o e^{-\mu z}\right) \tag{4.8}$$

Simplifying and dropping the logarithm produces

$$\frac{L_o}{L} = {I_o}^{\gamma} e^{-\mu z \gamma} \tag{4.9}$$

[4]Attenuation is used loosely in this section to refer to the product of the attenuation coefficient μ and the beam path length z.

[5]In diagnostic radiology conventional photographic film is used. Exposure to light causes the silver bromide grains of the film to become developable (to blacken).

Now L, the transmitted light through the film, is essentially the data which is digitised. Rearranging the equation above to make L the subject results in

$$L = L_o I_o^{-\gamma} e^{\mu z \gamma} \tag{4.10}$$

Equation 4.10 represents the raw digitised image. It is easily seen that the transformation which results in attenuation being additive is the natural logarithm. Making this transformation produces

$$\ln(L) = \ln(L_o I_o^{-\gamma}) + \mu z \gamma \tag{4.11}$$

which is an image where greyscale intensity is directly proportional to attenuation. This simple transformation, taking the natural logarithm of the image, is applied prior to any of the processing mentioned in this thesis.

Modelling the image of a nodule

Reconsider modelling the image of a nodule for the design of the template. We are interested in $nod(x, y)$, an analytical description of $I(x, y)$ (see equation 4.5) for the nodule. For a sphere of radius R, we have

$$z(x, y) = 2\sqrt{R^2 - r^2} \tag{4.12}$$

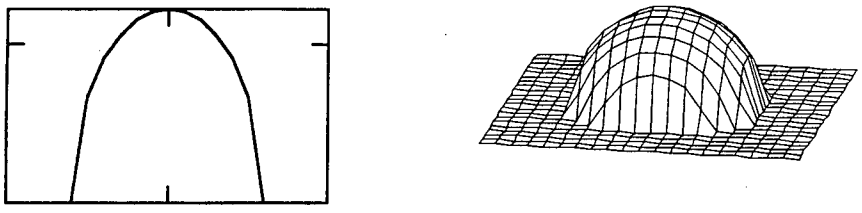
where r = displacement from the centre of the sphere. Substituting equation 4.12 into equation 4.5 gives $nod(x, y) = I_o e^{-2\mu\sqrt{R^2 - r^2}}$.

In conventional x-ray systems the image blurring is characterized by the point spread function (PSF). The x-ray beam is usually cone shaped and of fixed angular extent.

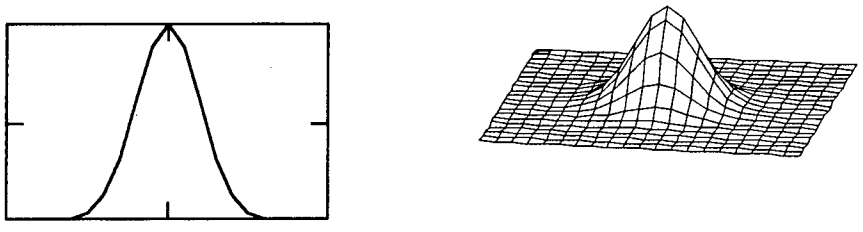
The PSF $beam(x, y)$ is convolved with the ideal nodule image $nod(x, y)$ to produce $nod'(x, y)$ as

$$nod'(x, y) = nod(x, y) * beam(x, y) \tag{4.13}$$

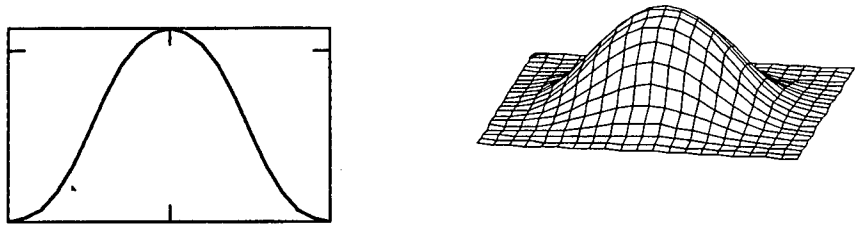
Illustrations of $nod(x, y)$, $beam(x, y)$ and $nod'(x, y)$ are provided in Figure 4.3.



Received x-ray energy $nod(x, y)$ after passing through a nodule



The point spread function $beam(x, y)$ is modelled here as a gaussian function



Convolving $nod(x, y)$ with $beam(x, y)$ yields $nod'(x, y)$, the image of the nodule

Figure 4.3: Depiction of the formation of a nodule image shown as profiles and surface plots. Note that the signal $nod(x, y)$ is shown inverted, as is typical of medical x-rays (which are photographic negatives).

The template model $nod'(x, y)$ can thus be constructed if R, μ, I_o, $beam(x, y)$, L_o and γ are known. These parameters could be determined empirically from the image by making several readings at the centre of, and at the background to, nodules in the image. Alternatively I_o and $beam(x, y)$ could be provided by the x-ray technician, μ could be read from a data book, L_o could be measured with a light meter, and γ is usually known. Another scheme for determining the template parameters is to compare the analytical model with a real nodule, extracted from the image[6]. The parameters of the analytical model can then be "tuned" until the best match is found.

[6] Any reference to *image* is a reference to the image which has been transformed into the linear attenuation domain, as it is this image on which all the processing is performed.

The model that is generated for the template matching procedure will be a function of the expected nodule radius R. It might be necessary to run the template matching procedure with a series of templates corresponding to different nodule radii.

As mentioned in the introduction to this chapter, there are factors which result in a loss of model accuracy.

Factors which reduce the accuracy of the nodule model

- **Irregular geometry**: Nodules are modelled as spherical objects. Real nodules are not perfectly spherical, and in many cases are considerably distorted. Also, the discrete nature of an image becomes significant for small circular objects.
- **Variable size**: Nodules of varying size often exist in the same lung.
- **Cavitation**: Certain types of tumours restrict the blood circulation at their centres. This results in the cells at the centre dying and thereby reducing x-ray attenuation for this region. This effect alters the image of the nodule and consequently decreases the accuracy of the model.

4.5 The Hough Transform as a Circle Finder

4.5.1 Introduction

The Hough transform is a method for detecting curves in images by exploiting the duality between points on a curve and parameters of that curve [2]. In particular, the Hough transform as a circle finder uses the information of the circle circumference only.

The Hough transform may be thought of as a template matching process which matches on abstracted features. According to Wechsler [32], "The Hough transform is conceptually nothing more than a match filter, which can be derived from the Radon transform and operated by accruing evidence pointing to some specific interpretation." The Hough transform was initially developed for straight line detection, and with this interpretation Wechsler's statement is valid. The Hough transform has been extended for the detection of arbitrary curves in images [2]. Although the straight line Hough transform is a special case of the Radon transform, generalisations of the Hough transform (for example, as a circle finder) are not instances of the Radon transform.

This discussion of the Hough transform, section 4.5, is divided into three sections. The first section reviews the concepts behind the implementation of the Hough transform as a circle finder, as proposed by Ballard *et al* [16]. The next section discusses the author's implementation, including modifications, of this algorithm. The final section briefly discusses Ballard's use of the Hough transform circle finder for the detection of cancer nodules in chest x-rays.

4.5.2 Overview of the Hough transform as a circle finder

The Hough transform is used primarily for line detection. In [16] this idea is extended to circle detection. The Hough transform has subsequently been generalised for the detection of arbitary shapes [2].

Ballard's Hough transform circle finder

The algorithm proposed by Ballard *et al* is essentially:

Start Process:

1. The original image is edge detected and thresholded to produce a binary image consisting primarily of lines and arcs. The binary image consists of black and white pixels only. The white pixels are those pixels which are the constituents of the lines and arcs.
2. Store the locations of all white pixels of the input image in a list.
3. Generate the accumulator array using the above list and the original image. The process of generating the accumulator array is explained below.
4. Threshold the accumulator array, as a function of radius, to find which elements of the array correspond to the centers of circles.

End Process.

The accumulator array is a three dimensional array. It may be written as $AA[r][x][y]$. Each of the $AA[r]$ is an image plane or a two dimensional layer of (x, y) coordinates. There are $R_{max} - R_{min}$ such layers, where R_{max} and R_{min} correspond to the maximum and minimum radius circles expected. The indicated number of layers assumes unit steps in r.

Let the input binary image to this algorithm be denoted $p(\underline{x})$, where $\underline{x}$ is the location coordinate (x, y). For each $\underline{x}$ there is a set $C_{\underline{x}}$ which contains the centre locations of all the circles

which could pass through the pixel at $\underline{x}$ for a radius R. $C_{\underline{x}}$ is thus the set of points on the circumference of the circle, centred at $\underline{x}$, having radius R. In Figure 4.4 the members of $C_{\underline{x}}$ are denoted by the shaded pixels. Each shaded pixel represents the centre of a circle which, with radius R, would pass through $\underline{x}$.

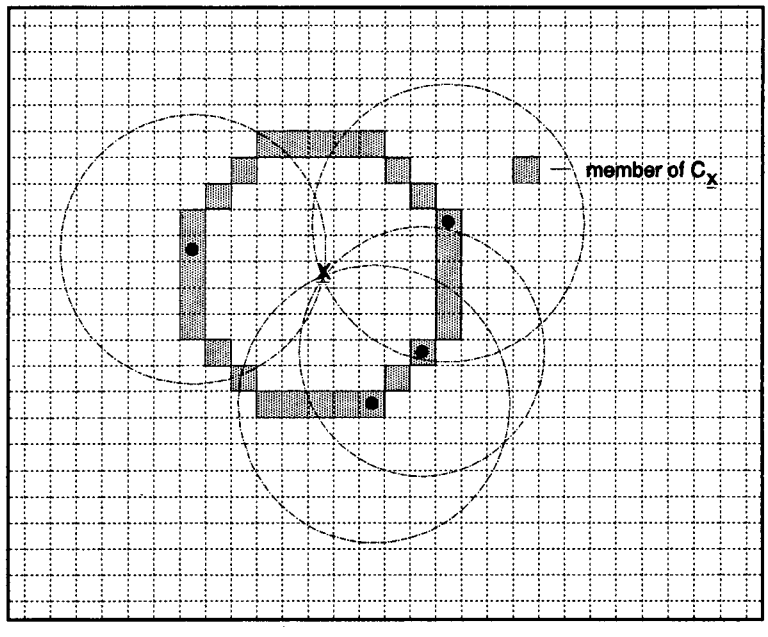


Figure 4.4: The set of locations of all circle centres which could pass through the current pixel, for a given radius R.

Let $\mathcal{X}_p$ denote the set of points $\{x \mid p(\underline{x}) = \text{white pixel}\}$. The Duda/Hart algorithm [9] is used for filling the accumulator array. $a(\underline{x})$ represents the circle's center location and $r(\underline{x})$ represents the circle's radius for each member of $C_{\underline{x}}$. The algorithm is shown below in pseudocode.

```
if x ∈ X_p then
      for all members of C_x
            find a(x)
            find r(x)
            increment AA at position AA[r][a]
      end for
end if
```

Ballard's optimizing techniques

Additional techniques were employed by Ballard *et al* to make the above algorithm computationally easier. These were:

1. A procedure for generating the members of $C_{\underline{x}}$ efficiently.

   To generate each member of $C_{\underline{x}}$, Ballard *et al* computed the digitization of $(r\cos\theta, r\sin\theta)$ for a fixed r, while θ increases in increments of $\Delta\theta$ ($\Delta\theta$ is a function of r).

2. Including the directional component of the gradient at the current location in the original image to eliminate unwanted members from $C_{\underline{x}}$.

   Instead of using all the points as members of $C_{\underline{x}}$, Ballard *et al* computed the direction of the gradient of the current pixel in the original image to select only members of $C_{\underline{x}}$ that are likely to be circle centers. This computation produces an angle ϕ for the direction of the gradient. Now only select the members of $C_{\underline{x}}$ that are in the range $[\phi - \frac{\Delta\phi}{2}, \phi + \frac{\Delta\phi}{2}]$. According to Ballard *et al*, this produces more accurate data and allows for faster computation of the algorithm.

4.5.3 Implementation of the Hough transform as a circle finder

The Hough transform has been implemented on a **SPARCstation 2** using Ballard's algorithm.

Discussion of Ballard's suggested computational improvements

- Ballard developed a procedure for generating the members of $C_{\underline{x}}$ efficiently.

  This procedure was implemented with the use of a sine and a cosine look up table. This addition resulted in a fourfold decrease in processing time. $\Delta\theta$ is computed as the reciprocal of the radius of the the nodule to be detected. The motivation for this decision is explained by Figure 4.5.

- Computational efficiency can be improved by including the directional component of the gradient of the pixel in the original image to eliminate unwanted members from $C_{\underline{x}}$.

[8]This minimum is a consequence of the discreteness of the circle which results in two adjacent members of $C_{\underline{x}}$ existing in a vertical or horizontal line. A maximum of $\sqrt{2}$ occurs where two adjacent members of $C_{\underline{x}}$ are adjacent diagonal pixels.

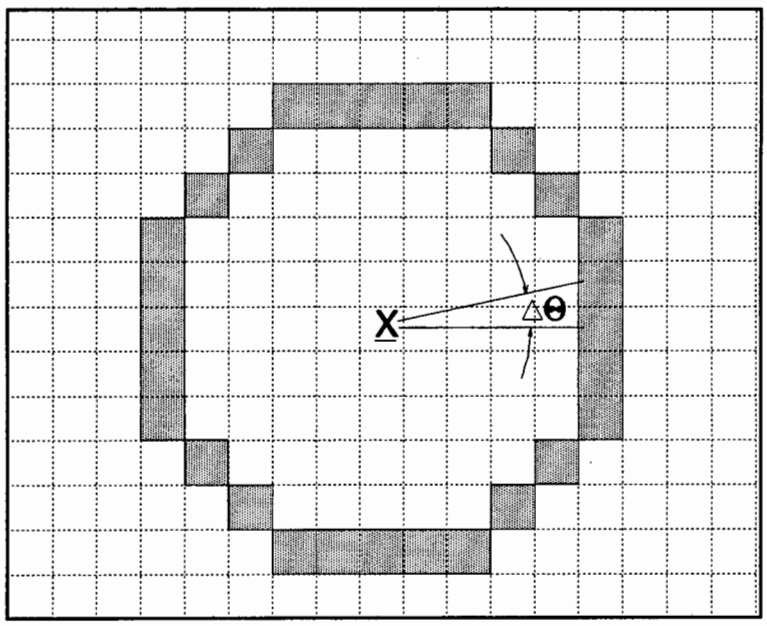


Figure 4.5: $\underline{x}$ represents the current circle center. The elements of $C_{\underline{x}}$ are the shaded pixels on the circumference of the circle centred at $\underline{x}$. The minimum angle $\Delta\theta$ is the angle subtended by two of the pixels of $C_{\underline{x}}$ to the circle centre $\underline{x}$. The distance from $\underline{x}$ to the pixels on the circumference is the circle radius R; and the distance between any two adjacent members of $C_{\underline{x}}$ is a minimum[8]of one. Trigonometry thus gives $\Delta\theta$ as $arctan(\frac{1}{R})$. For small $\frac{1}{R}$, we have $\frac{1}{R} \approx arctan(\frac{1}{R})$, and thus $\frac{1}{R}$ is a suitable equation for approximating $\Delta\theta$.

This idea was tested but is not being employed. The use of the directional gradient actually degraded the results.

The accumulator array, and classification

Normalization of the accumulator array is achieved by dividing the contents of each plane by the radius it corresponds to (this idea was also employed by Ballard *et al*). The motivation for this normalization is a consequence of the expected number of pixels in the circumference of a circle of radius R. This expected number of pixels is proportional to the radius by the relation: circumference $= 2\pi R$.

After normalization the weight of any element of the array may be directly compared with any other element. This is highly desirable because it allows a single threshold value to be used for the entire array. Deciding on a threshold value is, however, a burden. It presents an unwanted parameter. A similar problem is found in attempting to classify the correlation surface, which the template matching procedures generate. Techniques for dealing with the

Binary Hough Transform	Greyscale Hough Transform
Input image is a thresholded edge detected image	Input image is an edge detected greyscale image
Storage of the white pixels of the binary input image in an array is necessary	No such storage is required
The accumulator array elements in a circle around each white pixel are incremented by a constant	The accumulator array elements in a circle around every pixel is incremented by the intensity of the current center pixel

Table 4.2: Comparison of Binary and Greyscale Hough transform circle finding algorithms.

classification of both the accumulator array and the correlation surface are explored in chapter 5.

The greyscale Hough transform circle finder

The binary Hough transform, as proposed by Ballard *et al* requires two user input parameters. Any user input parameter is undesirable in an automated system and thus every effort is made to reduce the number of such parameters. The edge detected image must be thresholded to produce a binary image; this threshold value is the first parameter. The second parameter is the threshold value for the normalized accumulator array.

The threshold value for the edge detected image is very difficult to decide on without human intervention. To overcome this requirement a greyscale Hough transform was designed. This modification of the binary to the greyscale Hough transform is not a reversion to the Radon transform, as was explained in section 4.5.1.

The greyscale Hough transform is very similar in operation to the binary Hough transform. By way of explanation of the grey scale Hough transform, the differences between the binary and the greyscale Hough transform are supplied in Table 4.2.

The implications of a gray scale input to the Hough transform are far reaching:

- The burden of deciding on a threshold value, with or without human intervention, is eliminated.
- No information is lost by a thresholding process.

The binary Hough transform does, however, have some advantages over the greyscale trans-

form:

- It is considerably faster because only the white pixels need to be considered when generating the accumulator array. Contrast this to the greyscale transform which requires every pixel in the image to be considered.
- It is an established procedure with dedicated hardware and presumably optimized software available.

4.5.4 The application of the Hough transform as a circle finder to nodule detection in x-rays

Ballard *et al* used their algorithm for the detection of nodules in x-rays. They were successful in detecting large nodules (with radii, in pixels, of 6, 8 and 10) with no false hits. The success rate with smaller nodules was not as good. Edge detection was accomplished by filtering the image in the Fourier domain to enhance high spatial frequencies. The accumulator array threshold value was determined empirically. A threshold value, proportional to radius, was used in the accumulator planes corresponding to each radius in the search space.

4.6 Summary of the Image Analysis Modules

The analysis modules are summarised in terms of the required preprocessing and the requirements of invariance of match measurement and intensity matching (see Table 4.3).

From Table 4.3 we note that the variance match measure is the only technique which takes into account the intensity of the object being detected. The table also shows that a preprocessing technique is required, in all cases but the zero-mean template match, to ensure the correct operation of the analysis process.

Image analysis module	Required preprocessing for input image	Invariant measure of match	Intensity matching
Simple correlation measure	constant background level	NO	NO
Correlation measure with zero-mean template	none	NO	NO
Normalized correlation measure	constant background level	YES	NO
Normalized correlation measure with zero-mean template	constant background level	YES	NO
Variance match measure	constant background level	YES	YES
Binary Hough circle finder	edge detected and thresholded	YES	NO
Greyscale Hough circle finder	edge detected	YES	NO

Table 4.3: Comparison of image analysis modules.

Chapter 5

Classification and Evaluation

5.1 Introduction

The fourth and final phase of the generic automated nodule detection (A.N.D) system, proposed in section 1.2, is that of sub-image classification. According to the A.N.D model the purpose of this module is to classify sub-images, which correspond to suspected sites, in an image as either nodules or non-nodules.

Humans are considerably more adept at object recognition, and/or classification, than are computers. However, even humans find the task of classifying nodules in chest x-rays particularly difficult, which presupposes that the ability of computers to perform proficiently is minimal. An artificial intelligence (AI) algorithm is perhaps best suited to this classification task, as, in principle, an AI algorithm simulates the role of a human. The research of this thesis, however, explored into deterministic classification schemes only.

The two principle sections of this chapter are concerned with classification methods and with the evaluation of the automated nodule detection system.

5.2 Classification as a Postprocessor to Matching

The correlation surface is the output image produced by any of the template matching procedures. A single plane in the Hough accumulator array may also be considered a correlation surface. In a correlation surface, greyscale intensity is proportional to the degree of match.

Thus, the image analysis phase of the A.N.D system produces a surface, or surfaces, where peaks in the surface(s) correspond to the locations of suspected sites in the original image.

An intensity peak in the correlation surface implies that the object, in the original image, centred at this position is a good match with the template. When the match is not perfect, the energy is distributed in a small region around the true centre, as opposed to being a single pixel in the correlation surface. Thus, for imperfect matches, there are lower rounded hills instead of high intensity single spikes in the correlation surface.

The classification of sites by means of the correlation surface amounts to searching the correlation surface for high intensity values. Because of the likelihood of spreading at a single true centre, it is necessary to modify the nature of the search. Classification may be achieved by various techniques, such as (i) locating the X[1] most likely suspected sites, or (ii) locating all the suspected sites above a certain given threshold.

5.2.1 Classification using the Top X Suspects

This classification scheme requires a single parameter, *viz.* the requested number X of suspected sites. The following points describe an algorithm for determining the set of the X most likely nodule sites:

1. Each pixel of the image has the associated attributes of location and intensity. A list is generated of all the pixels.
2. This list is sorted[2] with respect to the intensity field of each pixel.
3. The first element of the sorted list (the coordinate of the pixel having the highest intensity) is accepted as the first element of the top X.
4. The i^{th} element of the sorted list is accepted as an element of the top X if it is not adjacent[3] to any of the previously selected top X elements. If it is adjacent to any

[1]The user specifies the number of requested sites.

[2]Conceptually, the list may be considered as being sorted. There are, however, computational techniques which can significantly speed up this process.

[3]This classification scheme assumes that individual, non-overlapping nodules are to be detected. Due to the imperfection of the match it is likely that the next highest intensity pixel is a neighbour of the formerly selected pixel. Therefore, without an adjacency test, small clusters of suspected sites would exist for a single nodule. It is only desired that each nodule is identified by a single suspected site at its centre. The adjacency test ensures that each suspected site is at least the distance of the sought nodule radius from any other suspected site.

of the selected elements then the $i + 1^{th}$ element of the sorted list is inspected. This continues until the top X non-adjacent elements are found.

5.2.2 Classification using a Threshold Value

This classification scheme requires a single parameter, *viz.* the threshold value above which all local maxima are to be classified. The following points describe an algorithm, based on a threshold value, for determining this set of nodule sites:

1. Each pixel of the image has the associated attributes of location and intensity. A list is generated of all the pixels which exceed the given threshold value.
2. This list is sorted with respect to the intensity field of each pixel.
3. The first element of the sorted list (the coordinate of the pixel having the highest intensity) is accepted.
4. The i^{th} element of the sorted list is accepted as an element if it is not adjacent to any of the previously selected elements. If it is adjacent to any selected element then the $i + 1^{th}$ element of the sorted list is inspected. This continues until all the non-adjacent elements of the sorted list are found.

5.2.3 A modification to classification if magnitude-only edge detection is used

When the Hough transform circle finder is used as the image analysis module, an additional requirement must be tested before accepting a suspect location as a hit. The suspected location must be examined to determine whether it is a hill (local maximum) or a valley (local minimum) in the original image. Only those sites which are hills in the original are accepted as hits. A valley, essentially a negative nodule, might give rise to a suspected site when using the Hough transform circle finder as a result of the edge detection preprocessing. All of the edge detectors mentioned, section 4.3, produced a magnitude-only output image. Consider a nodule described mathematically as $nod(x, y)$. Consider now an image which has a hill and a valley, both modelled as nodules at (x_1, y_1) and (x_2, y_2), superimposed on a constant level, DC. We can write the image as

$$Image(x, y) = DC + nod(x - x_1, y - y_1) - nod(x - x_2, y - y_2)$$

The magnitude-only edge detected image $EdgeIm$ would then be

$$EdgeIm(x,y) = EdgeNod(x - x_1, y - y_1) + EdgeNod(x - x_2, y - y_2)$$

where $EdgeNod(x,y)$ is the edge detected description of $nod(x,y)$.

Thus, if there are two structures in the image, which are identical but of opposite polarity, these structures will produce identical magnitude-only edge detected output. A solution to this problem is required in order to reduce the number of false hits. A simple modification is to compute the average greyscale intensity of a nodule sized region, in the original image, centred at the suspected site. The average greyscale intensity is then also computed for an annular region around the former area. False nodules, or valleys, are identified as having a higher average intensity in the surrounding annular region. Such sites are then excluded from the list of hits.

5.2.4 Additional classification ideas

Circularity in the correlation surface

Correlation of a circular object with a circular template yields a peak in the correlation surface which is circular in cross section. Accordingly, in order to discriminate between nodule and non-nodule structures Giger *et al* used a circularity measure on the correlation surface [12]. The correlation surface was then thresholded at consecutive levels and the circularity of the "islands" in the thresholded image was graphed. It was expected that peaks corresponding to matched nodules would exhibit similar circularity for a range of thresholds. The effective diameter of a thresholded island was computed (see Figure 5.1) as

$$2\sqrt{\frac{area\ of\ thresholded\ island}{\pi}}$$

They then computed the degree of circularity, of each island, as the ratio

$$\frac{area\ of\ island\ within\ circle}{area\ of\ island}$$

It was anticipated that peaks corresponding to nodules would exhibit both a high degree of circularity and a fairly constant circularity measure over the range of thresholds. Peaks which did not exhibit this tendency were less likely to correspond to nodules.

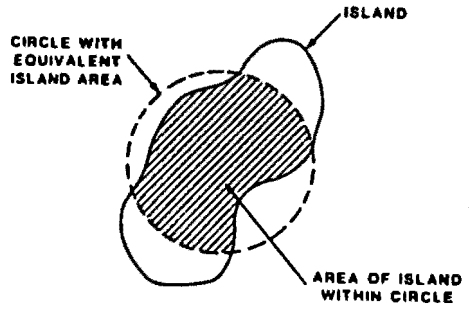


Figure 5.1: Giger's circularity measure

Use of multiscale measures

The Hough transform as a circle finder generates an accumulator array of several planes, with each plane corresponding to a different radius. This array can be inspected for peaks within a three dimensional search space. The accumulator array has the advantage of being size invariant. A similar effect can be achieved by producing a range of correlation surfaces, each generated by a template of increasing radius.

Multiscale measure can be used to examine the behaviour of peaks in the correlation surface as the radius[4] of the template is varied. In this way, knowledge of the multiscale behaviour of peaks corresponding to nodules can be utilised for additional discrimination.

The converging squares algorithm

The converging squares algorithm [21] is not well suited to peak detection in the correlation surface because the ideal match in the correlation surface is an intensity spike, whereas the converging squares algorithm looks for high density[5] regions.

[4]In fact, any variable which modifies the correlation surface may be used to aid in differentiating between nodule and non-nodule peaks in the correlation surface. For example, the response of a peak in the correlation surface when a parameter of the preprocessor is varied but the template size is kept constant may produce a distinctive trend which could aid in discrimination.

[5]Density is defined as the sum of the greyscale intensities, within a region, divided by the pixel area of that region.

5.3 Evaluation of the Automated Nodule Detection System

In order to evaluate the automated nodule detection (A.N.D) system it is necessary to quantify its success rate. Such a method which quantifies the success of the A.N.D system also allows the relative suitability of each of the modules and their parameters to be evaluated. In this way the A.N.D system can be constructed from the combination of enhancement and analysis modules which provide the highest success rate.

A quantifiable method of evaluating the automated nodule detection system is the receiver operating characteristic (ROC) plot.

5.3.1 The receiver operating characteristic

The receiver operating characteristic (ROC) is a standard measure in detection theory. In most detection processes there is a parameter, often a threshold value, which, when varied, determines the probability of a false hit P_f and the probability of a true detection P_d. This locus of probability pairs is known as the ROC plot [10].

P_d is defined as the ratio of the number of true hits to the number of true sites. A hit is a location returned by the detection process as a suspected site. The true sites are known *a priori* and the hits are counted against the true sites to determine the number of true hits. A true hit corresponds to an element of the set $\mathcal{C} \cap \mathcal{D}$ in Figure 1.1 ($\mathcal{A}$ represents the set of all pixels in the image, $\mathcal{C}$ represents the set of pixels classified by the computer as nodules, $\mathcal{D}$ represents the set of pixels corresponding to true nodule sites).

P_f is defined as the ratio of the number of false hits to the number of false sites.[6]

In Figure 1.1 a false hit is an element of the set $\mathcal{A} \cap \overline{\mathcal{D}}$.

---

[6] For the case of non-overlapping nodules, the total number of sites in the image is:

$$Total_{sites} = \lceil \frac{M}{vertical\ tolerance} \rceil . \lceil \frac{N}{horizontal\ tolerance} \rceil$$

where M is the number of rows and N is the number of columns in the image. The vertical and horizontal tolerances refer to the error distances from the true site within which a hit is regarded as being true. These tolerances are usually taken as the vertical and horizontal radii of the nodule, respectively. In general, the vertical and horizontal nodule radii are equal but certain image acquisition systems might blur more vertically, or vice versa.

The number of false sites is given as:

$$False_{sites} = Total_{sites} - True_{sites}$$

The line $P_d = P_f$ implies equal probability of true and false-positive, and is thus indicative of no improvement in detection. The generic ROC plot showing the line $P_d = P_f$ is illustrated by Figure 5.2.

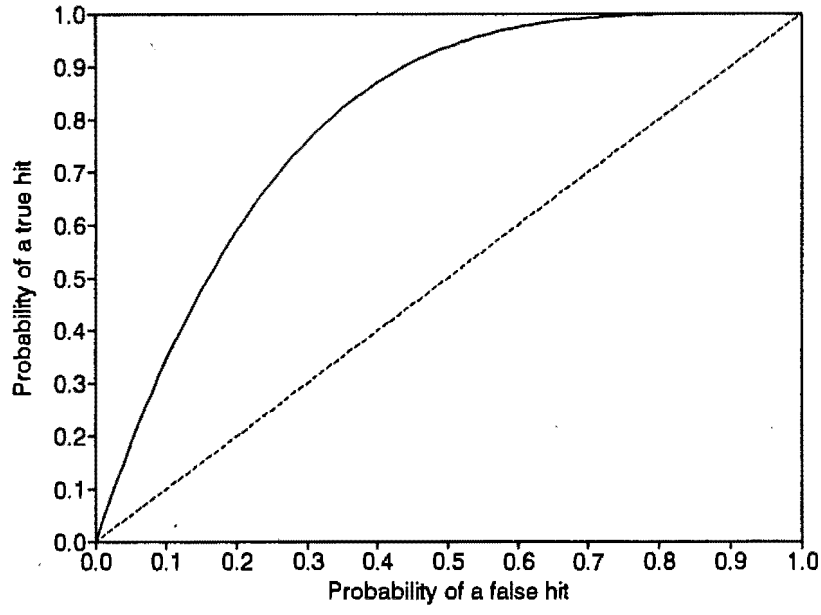


Figure 5.2: Generic receiver operating characteristic (ROC). The dashed line represents $P_d = P_f$, and the solid curve is a typical ROC.

In some cases, the horizontal axis is plotted using a logarithmic scale as the number of false sites may be several orders of magnitude greater than the number of true sites. This results in the line $P_d = P_f$ appearing as an exponential curve.

The ROC plot can be used to evaluate enhancement techniques, detection techniques and even the accuracy of the model used as the template in template matching.

5.3.2 Generation of the ROC plot

The following points describe the procedure for producing the ROC plot:

1. Generate a list of the coordinates of the true sites by reading from a file containing *a priori* coordinate data.
2. Classify the correlation surface for a range of thresholds, using the classification scheme of section 5.2.2. This returns a list of the coordinates of the suspected sites.

3. Inspect each suspect and determine whether it is close enough[7] to a true site to be counted as a true hit. Those that are not close enough are counted as false hits. Thus for each threshold value a percentage of true hits and a percentage of false hits is determined.

4. Accumulate the graph data using the percentage of true hits and percentage of false hits for each threshold value.

It is difficult to extract a single number representing the success rate from the ROC plot. Numerous success measures[8] are possible, but none of these measures are truly representative of the full graph.

An automated system for nodule detection is intended for use as a routine screening procedure in a clinical environment. The objective of such a system is to alert a diagnostician of the possibility of nodules in an x-ray. Only the medical cases associated with these x-rays would be subjected to careful further examination. Therefore, it is vitally important that all the x-rays containing nodules are isolated as suspect. On the other hand, it is also important that the number of false diagnoses be kept to a minimum. It would be disconcerting and expensive (in both time and follow-up procedures) to subject clear x-rays to further examination. It is apparent from the ROC plot that ensuring a high probability of detection also ensures a large percentage of false alarms, and therefore a compromise must be made between the certainty of diagnosis and the allowance for false alarms.

The percentage of true hits for a constant false alarm rate is a reasonable measure of success. However, it is not a robust measure because ROC plots of substantially different shapes might give the same reading at this single point, or vice versa. To increase the reliability of the measure it was decided to use the area under the *logarithmic* ROC plot between the 1% and 10% false hit rate points (see Figure 5.3). The range, 1% to 10% false hits, was taken as the compromise between detection certainty and a low false alarm rate. The maximum success measure, or the maximum area, is the area of the rectangle having a width in the range -1 to -2 and a height of 100%. In general, shifting the ROC plot up and to the left increases the success measure.

The ROC plot gives an automated measure of the accuracy of the A.N.D system. It cannot, however, be used in the actual A.N.D system since its generation and measurements require

[7]The decision of *close enough* is related to the tolerance region discussed in section 5.3.1. At present, a suspect is regarded as being close enough if it is within the nodule radius of the true site coordinate.

[8]The success measures which have been considered are (i) the area under the plot, (ii) the percentage of true hits when the false hit percentage is 0.1%, 1% or 10%, (iii) the maximum vertical difference between the ROC plot and the 50% curve.

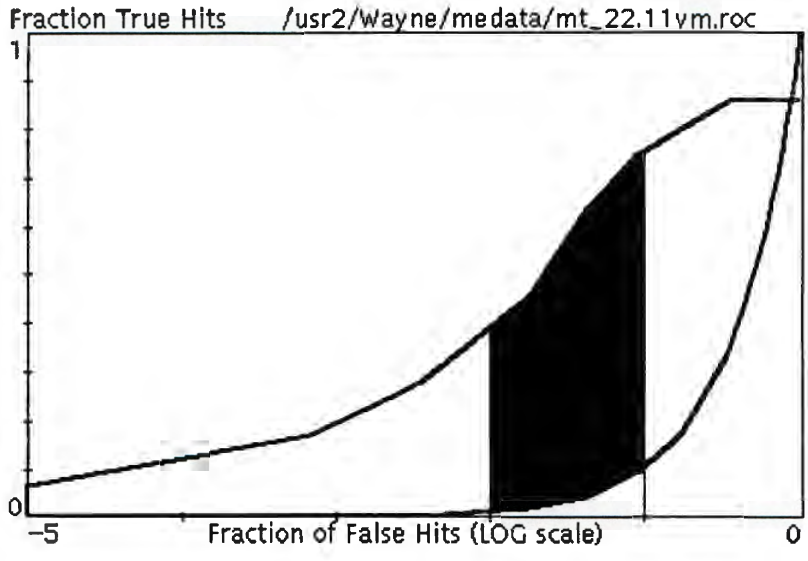


Figure 5.3: An ROC plot showing the area (shaded) between the 10% (-1) and 1% (-2) false hit rates. The area, in the same range, under the exponential curve is a constant for all ROC plots.

prior knowledge of the true sites in the image. A particular threshold, or a requested number of suspected sites, or a combination of both, must still be chosen for the implementation of the A.N.D system.

Chapter 6

Results

6.1 Introduction

In this chapter, the effect of many of the techniques discussed in previous chapters, as applied to a medical chest x-ray is illustrated, and a quantitative analysis of the most suitable enhancement and analysis modules is presented. There are five main sections: (i) the development of a test image, (ii) the results of different enhancement techniques as applied to the test image, (iii) the results of different analysis techniques as applied to the enhanced test image, (iv) an evaluation of the most suitable enhancement module, and (v) an evaluation of the most suitable analysis module.

6.2 Development and Description of the Test Image

A medical chest x-ray was obtained from the film library in the Department of Radiology, University of Cape Town. With the assistance of Dr Stoner [28], a specialist radiologist, this particular x-ray film was chosen, as it showed several indistinct metastatic[1] nodules. Dr Stoner then diagnosed the x-ray and indicated the locations of the nodules.

The x-ray film was then digitised using an **OmniMedia 6cx XRS** scanner. The digitised image had a bit depth of 8 and a resolution of 2400 rows by 3000 columns. The image was

[1] A metastatic nodule is a secondary cancer growth. Lung metastases are generally more spherical than primary lung cancer nodules.

then reduced in resolution and improved in bit depth with the use of superpixel binning[2]. This produced a 12 bit image of 600 rows by 750 columns.

The scanned x-ray image is illustrated in Figure 6.1. The image is then transformed into the linear attenuation domain (see section 4.4.6), as shown in Figure 6.2.

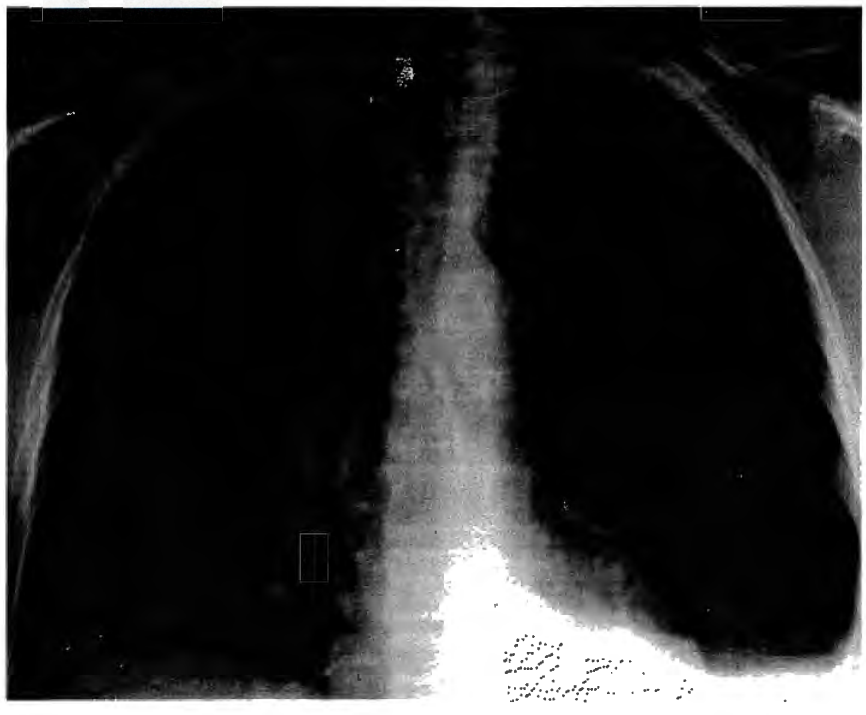


Figure 6.1: The scanned x-ray image from which the test image is derived.

The x-ray image contained only four nodules of a similar size. Since the success measure is statistical in nature, it was important that many specimens existed in the image. The best nodule specimen was extracted from the image and an analytical model was "tuned" until it best matched the real nodule, as discussed in section 4.4.6. A radius of 8 pixels for the nodule model was found to be the best match with the real nodule. 141 instances of this analytical model were then superimposed on the image[3] in a regular fashion within the lung region of the x-ray. A close-up view and a surface plot of a region of the image, illustrating the real nodule and a simulated nodule, are shown in Figures 6.3 and 6.4, respectively. The image of the simulated nodule, shown alongside the real nodule, is largely corrupted by the structured noise of the picture. Being superimposed on a relatively noiseless background, the image of

[2]The image was effectively tesselated into subregions of 4x4. A superpixel was then produced as the sum of the pixels in each subregion.

[3]Such a linear superposition is only valid because the image has already been transformed into the linear attenuation domain (see section 4.4.6).

Figure 6.2: The x-ray image transformed into the linear attenuation domain.

another simulated nodule (in Figure 6.5) is more easily discerned.

The test image with the superimposed nodules is depicted by Figure 6.6. It is very difficult to see where the simulated nodules are in the test image. For this reason, an enhanced version of the test image is included for the sole purpose of allowing the reader to see the locations of the nodules (see Figure 6.7) .

All processing was done on a **SPARCstation 2** running custom developed C code on the SunOS 4.1.2 operating system. The image viewing utility and the surface plot utility was provided by **Khoros** [7]. The relevant C source code is described in Appendix A.

6.3 Results of the Enhancement Techniques

The effect of unsharp masking (see section 2.3.1) applied to the test image is illustrated. An example of the result of linear unsharp masking is shown in Figure 6.8. Note the shadow band artifacts which are particularly evident at the ribs. An example of median unsharp masking is provided in Figure 6.9. The triangular artifact of symmetrical two dimensional median filtering is not obvious in the image.

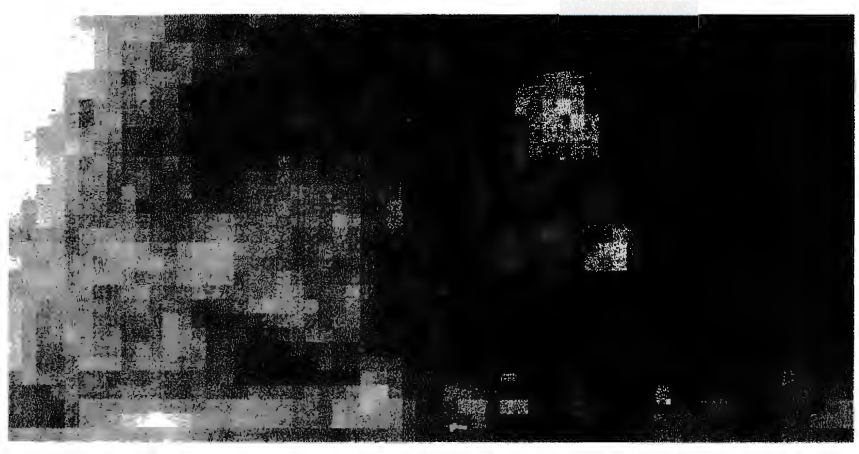


Figure 6.3: A close-up view of a real nodule (left) and a simulated nodule (right) in the test image.

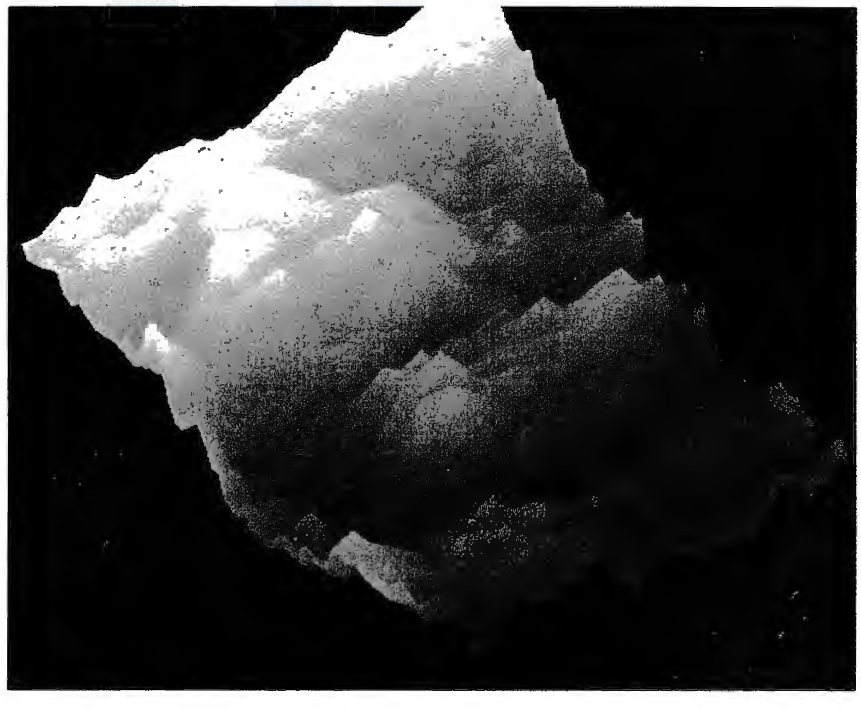


Figure 6.4: Surface plot of the real nodule (above left) and the simulated nodule (below right) in the test image.

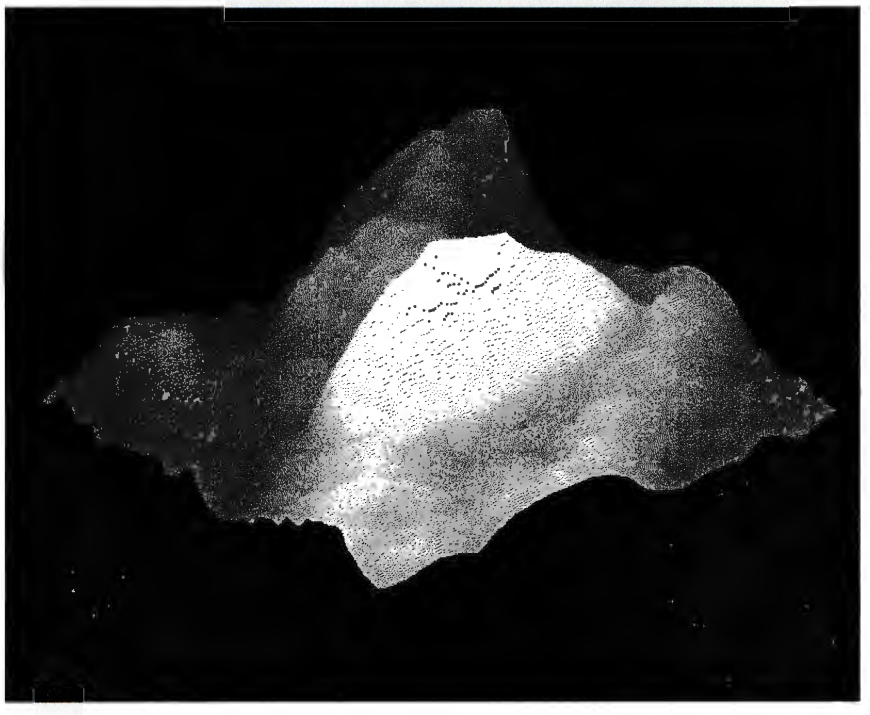


Figure 6.5: Surface plot of an easily discernible simulated nodule in the test image (the perspective of the surface plot causes the nodule to appear elongated).

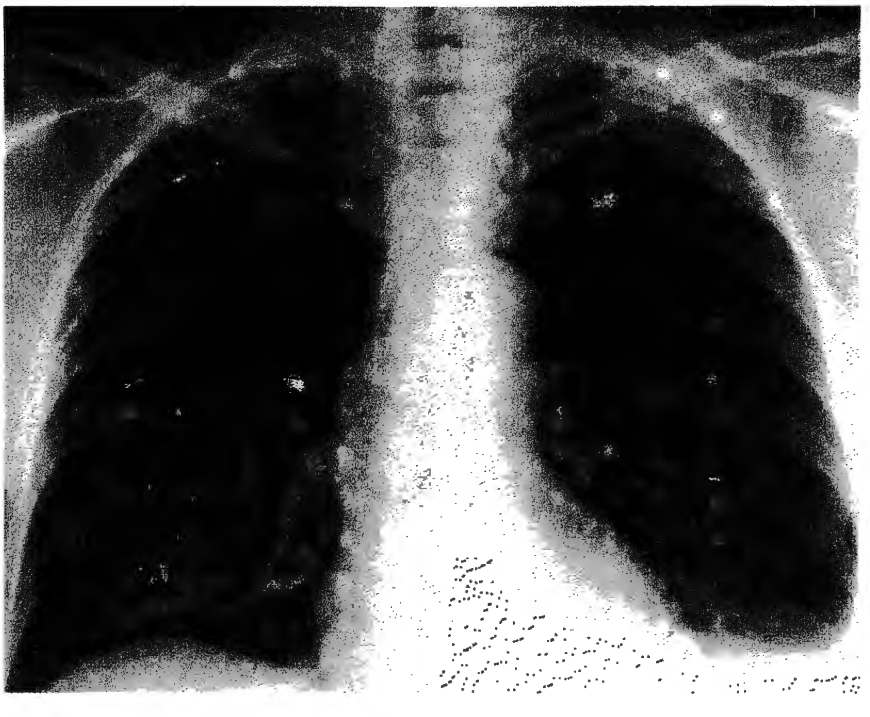


Figure 6.6: The test image with superimposed nodules.

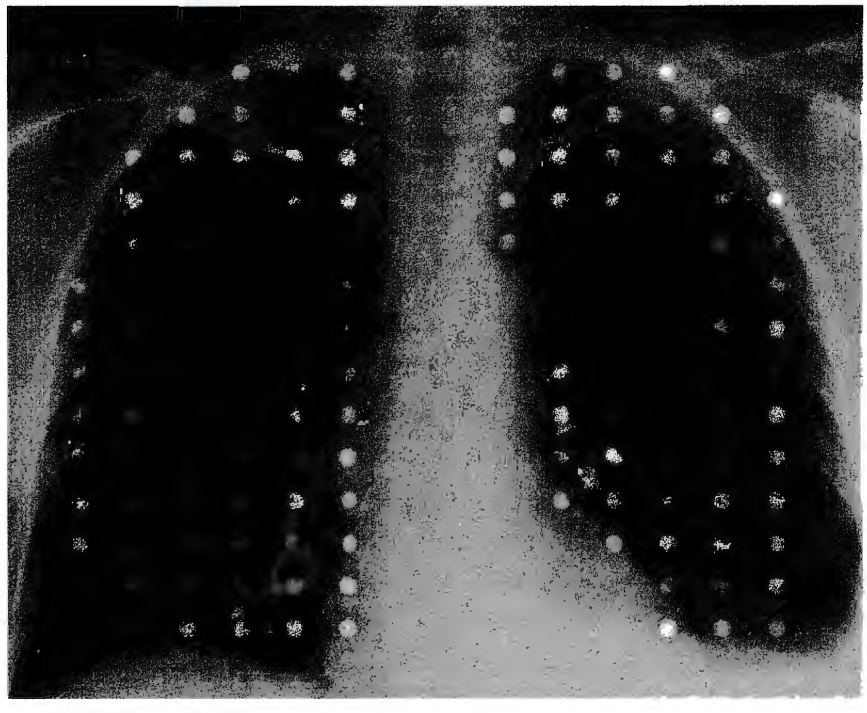


Figure 6.7: An enhanced version of the test image illustrating the locations of the nodules.

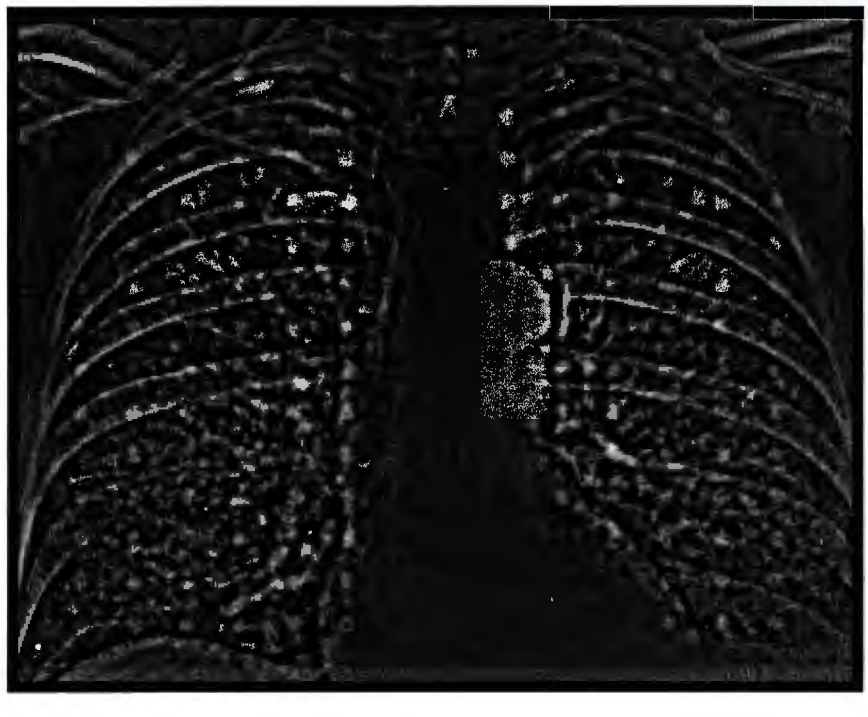


Figure 6.8: Linear unsharp masking with a circular footprint inscribed in a 21x21 footprint.

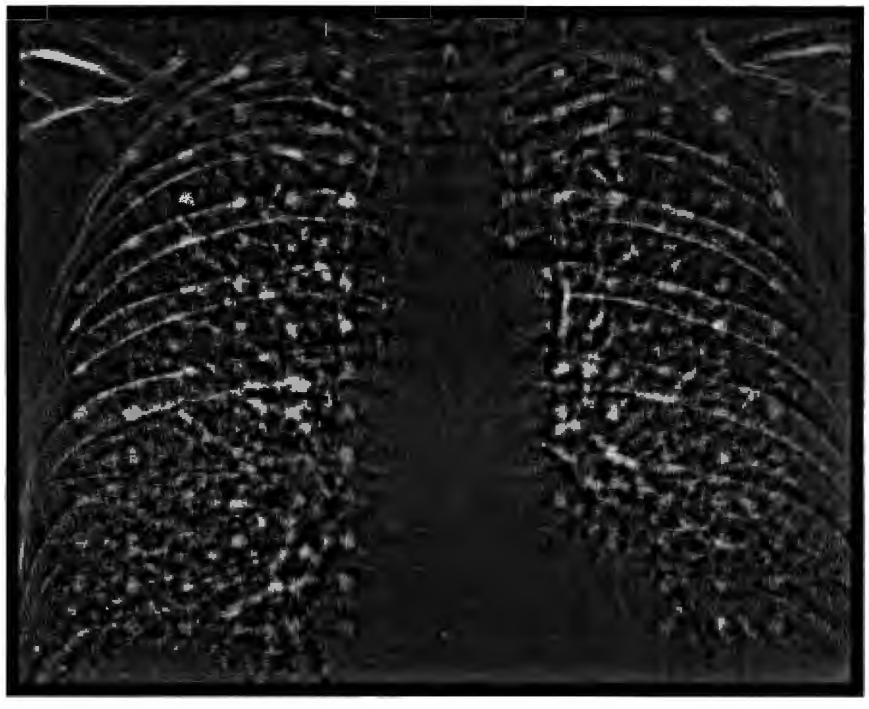


Figure 6.9: Median unsharp masking with a circular footprint of radius 11.

Examples of statistical difference filtering (see section 2.3.2) are illustrated in Figures 6.10 and 6.11. Shadow band artifacts are prominent and are proportional in width to the footprint of the filter. There is also an over-enhancement of noise, particularly when a smaller footprint is used (see Figure 6.10).

The effect of adaptive histogram equalisation (see section 2.3.3), applied to the test image, is shown in Figure 6.12. The over-enhancement of noise is clear.

The result of the rotating rod algorithm (see section 3.3) applied to the test image is provided in Figure 6.13. It is clear from the figure that the rotating rod has the effect of eliminating all image data that is not smaller in all orientations than the length of the rod, which is 20 pixels in this example (see section 3.6).

The greyscale top-hat transform (see section 3.5) involves subtracting an opened image from the original image. The size and type of the structuring element (SE) employed dictates the nature of the enhancement. Figure 6.14 is an example of the test image top-hat transformed using a disc SE of radius 8, while Figure 6.15 shows the result of a sphere SE of radius 10. These figures (6.14 and 6.15) have been processed with SEs of similar footprint dimensions to the rotating rod, but, as is noticeable from the figures, all structures which are smaller in any orientation than the footprint size (17x17 or 21x21) are retained. This differs from

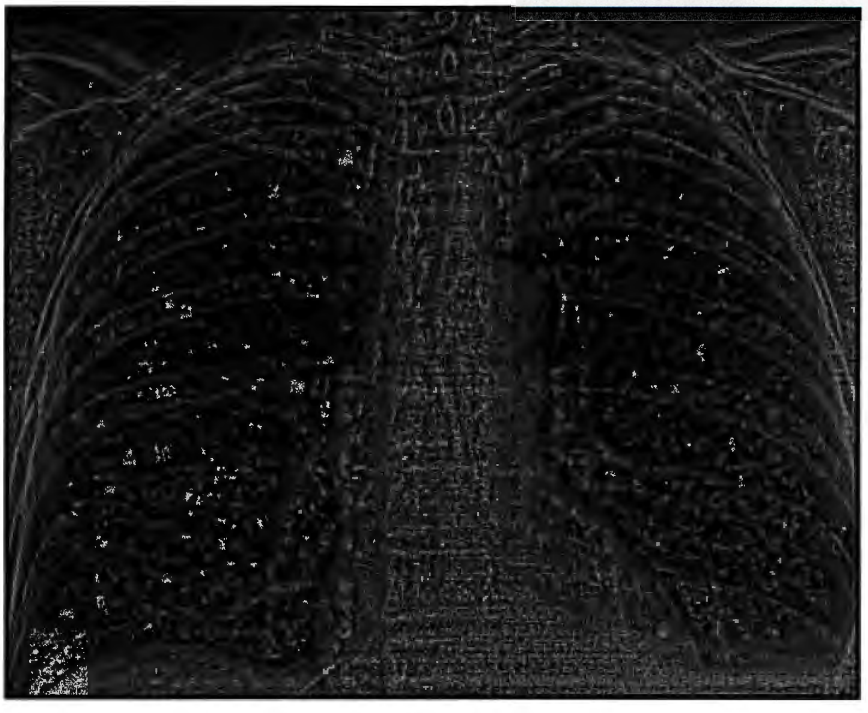


Figure 6.10: Statistical difference filtering with a footprint of dimension 9x9.

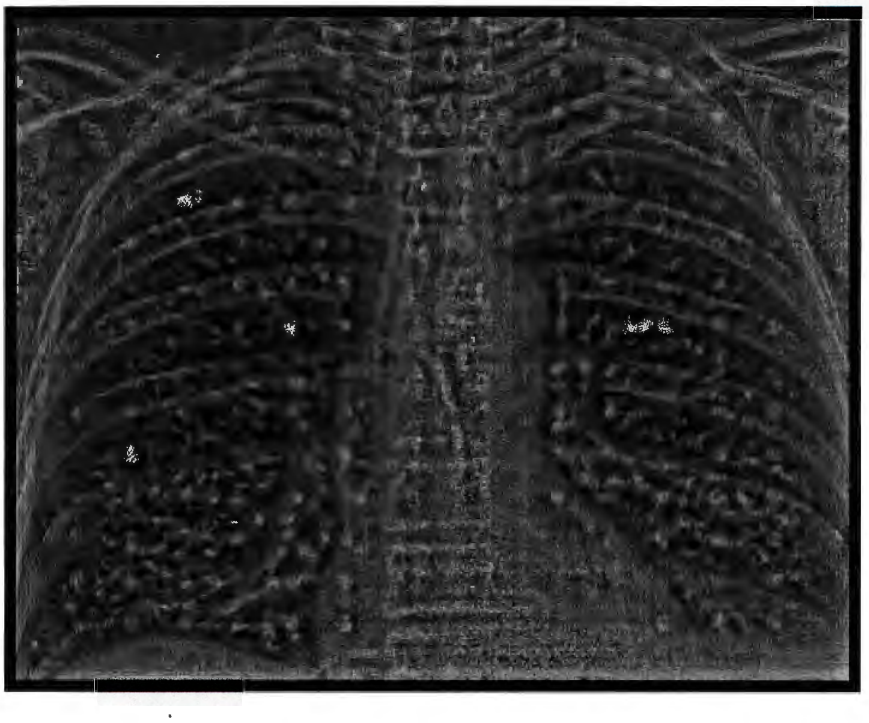


Figure 6.11: Statistical difference filtering with a footprint of dimension 21x21.

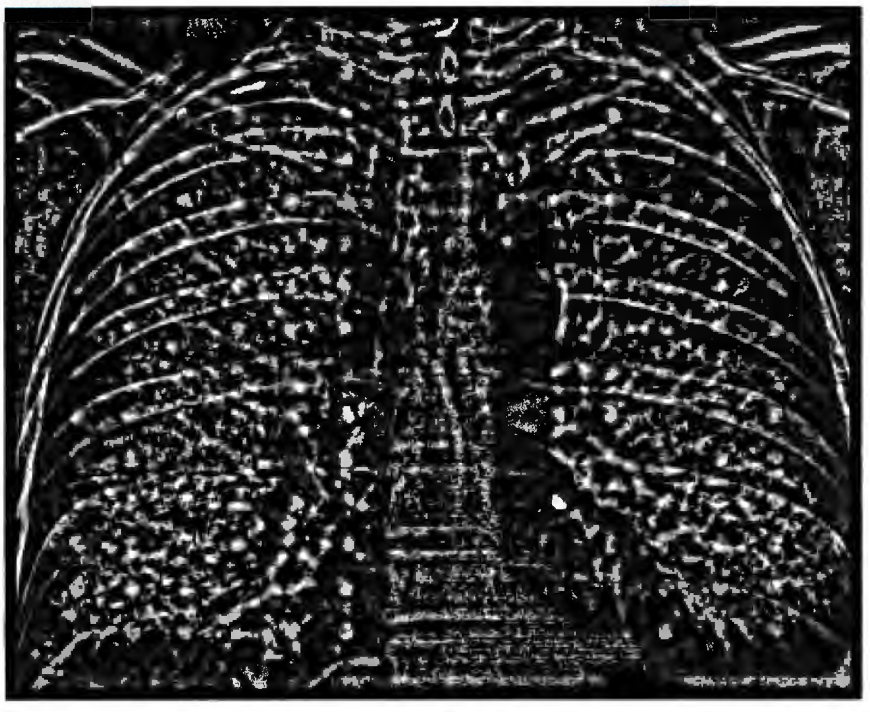


Figure 6.12: Adaptive histogram equalisation with a footprint of dimension 21x21.

Figure 6.13: Rotating rod algorithm with a rod length of 20 applied to the test image.

the rotating rod where almost all elongated image data is removed (see section 3.6 for an explanation of this phenomenon). The small amount of remaining rib structure in Figure 6.13 could, in principle, also be eliminated if the rotating rod was implemented to operate at more orientations (than the current eight).

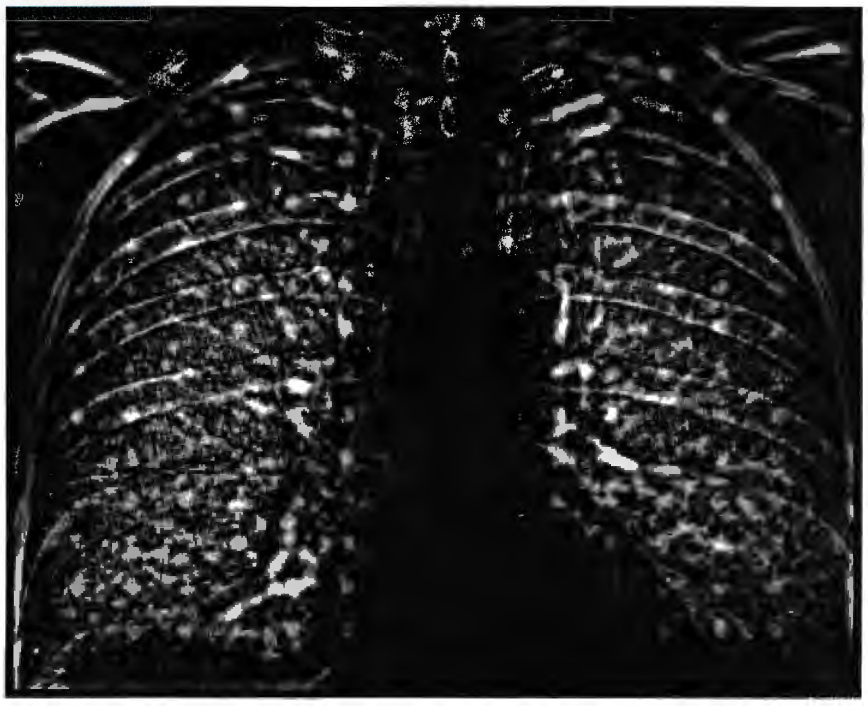


Figure 6.14: Test image top-hat transformed with a disc SE of radius 8.

6.4 Results of the Analysis Techniques

The results of the different edge detectors (see section 4.3) applied to the test image are shown in Figures 6.16, 6.17 and 6.18. Sensitivity to noise is apparent in the Sobel edge detected image, Figure 6.16. The difference of areas edge detected image, Figure 6.17, is less sensitive to noise but the edges are blurred. The nonspread edge detected image, Figure 6.18, shows the relative insensitivity to noise and crisp edges.

In a later section (6.5) it is shown that the image enhanced with median unsharp masking is well suited to the subsequent analysis procedure of variance matching. This enhanced image, already shown in Figure 6.9, is used as the input image to the variance matching procedure. The correlation surface resulting from variance matching the preprocessed image is provided in Figure 6.19. The same input image is subjected to normalized correlation template matching,

Figure 6.15: Test image transformed with a sphere SE of radius 10.

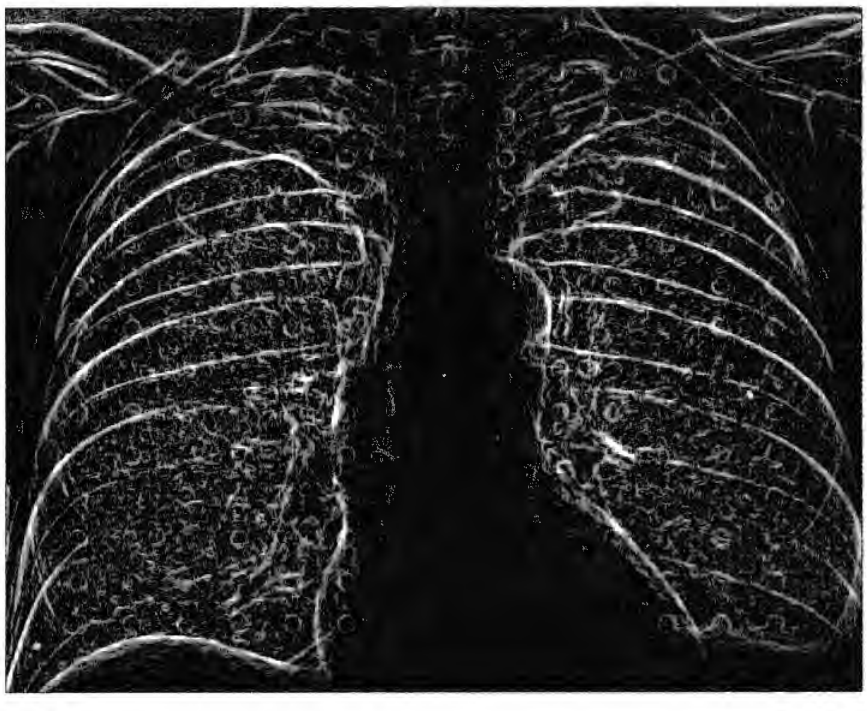


Figure 6.16: Edge detected test image using the sobel edge detector.

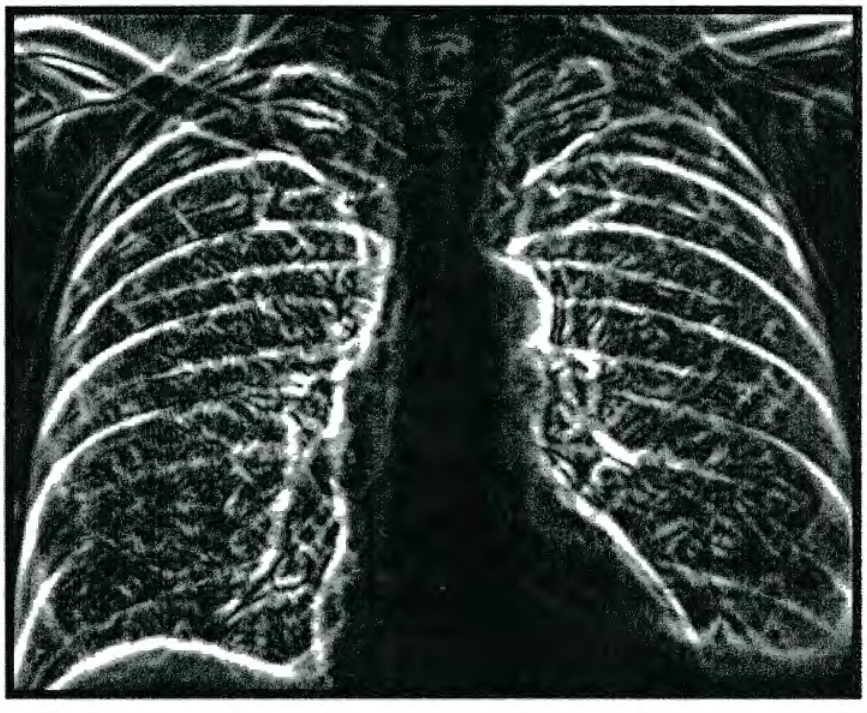


Figure 6.17: Edge detected test image using the difference of areas edge detector.

Figure 6.18: Edge detected test image using the nonspread edge detector.

and the resultant correlation surface is shown in Figure 6.20.

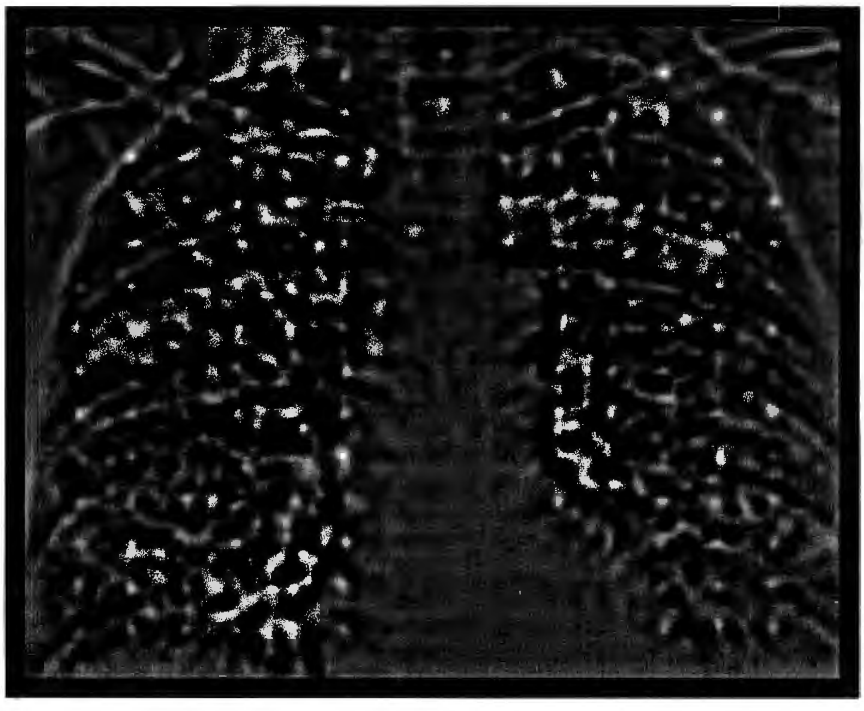


Figure 6.19: Correlation surface produced by variance matching with the median unsharp masked image as input.

The Hough transform as a circle finder generates an accumulator array, a single plane of which is analogous to a correlation surface. Figure 6.22 depicts a plane of the accumulator array of the binary circle finder, while Figure 6.23 shows a plane of the accumulator array of the greyscale circle finder. Both accumulator arrays were generated after edge detecting the test image with the nonspread edge detector. Naturally, the edge detected image, which is the input to the binary circle finder, was also thresholded (see Figure 6.21).

6.5 Evaluation of the most suitable Enhancement Module

Chapters 2 and 3 investigated different techniques for the implementation of the A.N.D enhancement module (see section 1.2). This section discusses the evaluation of various of these enhancement techniques and their parameters. The measure of success of an enhancement process in the A.N.D system is gauged by the area of the ROC plot between the false hit rates of 10% and 1% (see section 5.3.1 and Figure 5.3).

The objective of this experiment is twofold. Firstly, the enhancement module which is the

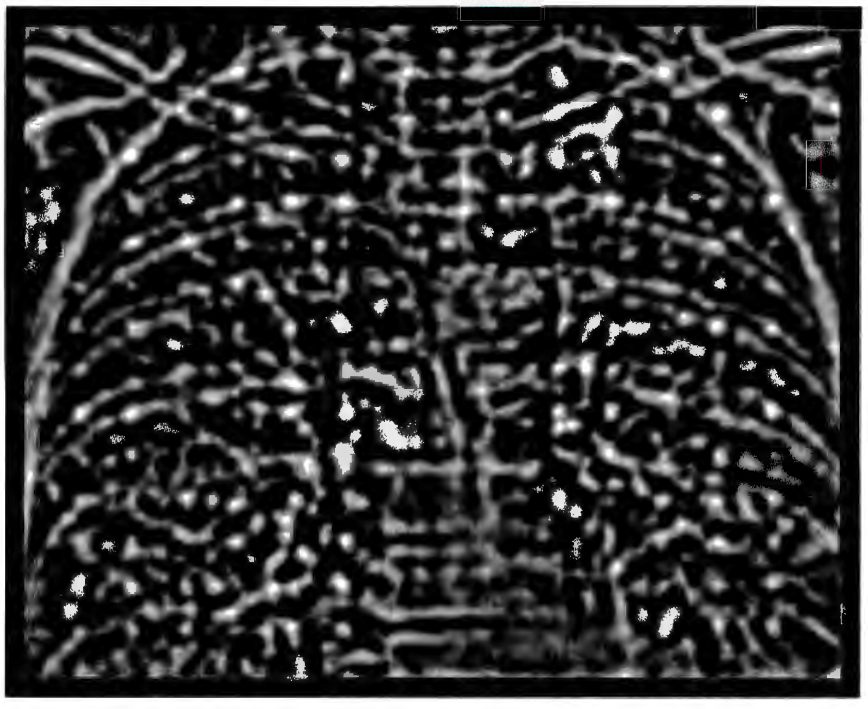


Figure 6.20: Correlation surface produced by normalized correlation template matching with the median unsharp masked image as input.

most suitable as the preprocessing stage to automatic detection is to be determined. Secondly, the parameters of this enhancement module, with respect to nodule diameter, are to be established.

The experiment was conducted by subjecting the test image (see section 6.2) to an enhancement process followed by variance matching. The template for variance matching is generated by subjecting a simulated nodule to the same preprocessing which produced the enhanced image from the test image. The correlation surface, thus produced, is classified using the threshold value method (see section 5.2.2) for a number of thresholds, to produce an ROC plot. The success measure is computed from the data in the ROC plot. A maximum measure of 100 area units is possible (see section 5.3.2). This process is repeated for a range of parameters within each enhancement technique. The success rate versus an enhancement parameter (often footprint size) is graphed for each of the enhancement implementations (see Figure 6.24). The dimension of a footprint is often referred to as radius in the following sections. In this context, a footprint of 13x13 pixels is said to have a radius of 6, ($\lfloor \frac{13}{2} \rfloor$).

Note that this success measure does not refer to the success of the particular enhancement technique, but to the success of the combination of that enhancement technique with a vari-

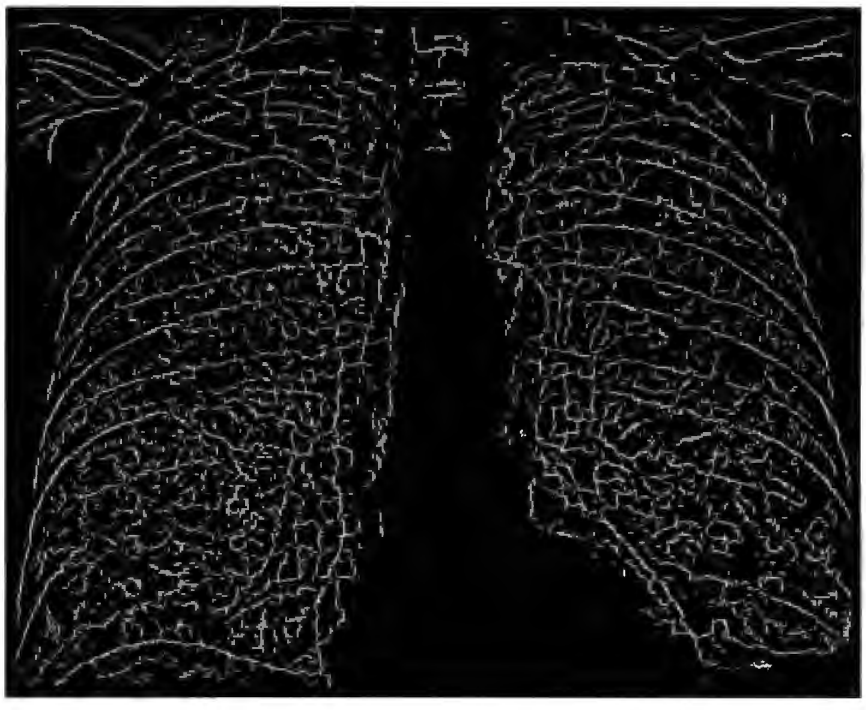


Figure 6.21: Thresholded edge detected test image which is used as the input to the binary Hough transform circle finder.

ance matching detection stage. However, because all of the enhancement techniques are followed by an identical detection process, the success measure allows a relative comparison of the enhancement techniques as preprocessors to variance matching. It is possible that an enhancement module which produces a poor success measure when used as a preprocessor to variance matching, will produce a good success measure when used as the preprocessing stage to a different detection stage, and vice versa.

The ROC plot used for determining the success measure was generated using ten equally spaced threshold values. A plot consisting of only ten points is often very coarse and sometimes quite inaccurate. Using more points to generate a smooth, reliable ROC plot is not difficult, but it is a considerable computational burden. The immense system requirements made exhaustive testing of the various enhancement and analysis modules impossible. In principle, these tests which attempt to evaluate the most suitable modules and their parameters, for the A.N.D system, should use all possible combinations of the modules and their parameters. The tests for the enhancement techniques alone required approximately 75 hours of CPU time and more than 500 Megabytes of disk space.

The following sections evaluate the suitability of unsharp masking, median unsharp masking

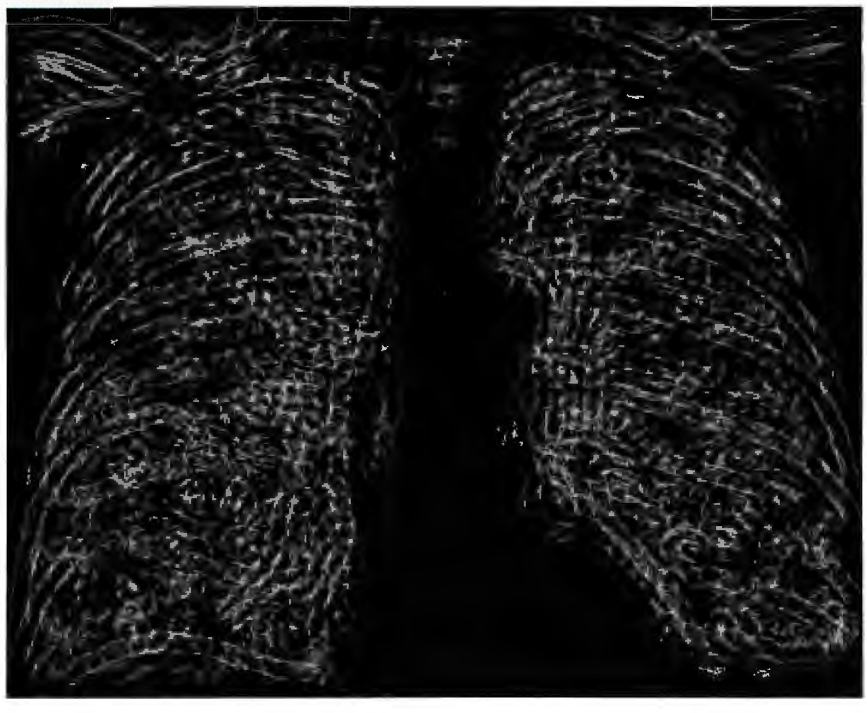


Figure 6.22: A plane of the accumulator array, corresponding to a radius of 8, of the binary circle finder.

and the top-hat transform as enhancement techniques. These processes were chosen because they are the enhancement techniques which satisfy additive invariance, a requirement of the variance matching stage.

6.5.1 Suitability of linear and nonlinear high pass filtering to automatic nodule detection

An experiment to determine the optimum footprint size of the linear filter used in the unsharp masking process was conducted. The linear filter used had a circular footprint and was a simple[4] averaging finite impulse response filter.

The graph of the success rate versus the footprint radius is provided in Figure 6.25.

This graph indicates that the combination of unsharp masking and variance matching is very poor for the entire range of filter radii. A notable feature of the graph is the improved performance for small filter radii. A small filter radius implies that the image is, in essence,

[4]A windowing function was not used.

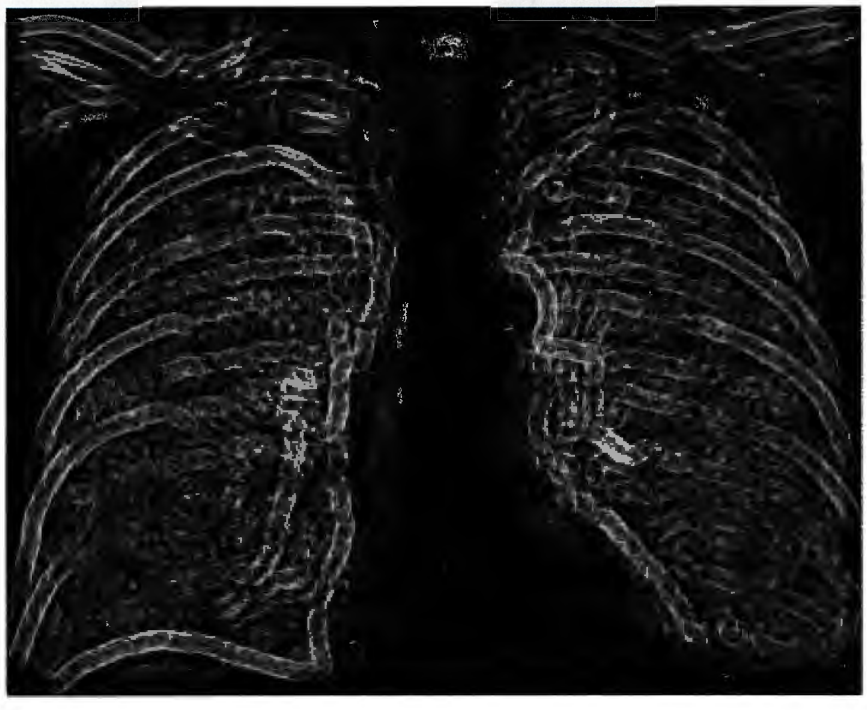


Figure 6.23: A plane of the accumulator array, corresponding to a radius of 8, of the greyscale circle finder.

being edge detected. This is in accordance with the opinions of the authors mentioned in section 4.4.5.

An experiment to determine the optimum footprint size of the median filter used in the median unsharp masking process was conducted. This experiment was performed for the square footprint and the circular footprint median filter. It was also done for the one dimensional dual pass median filter (see section 2.3.1). The graph of success rate versus footprint radius for each median filter type is supplied in Figure 6.26.

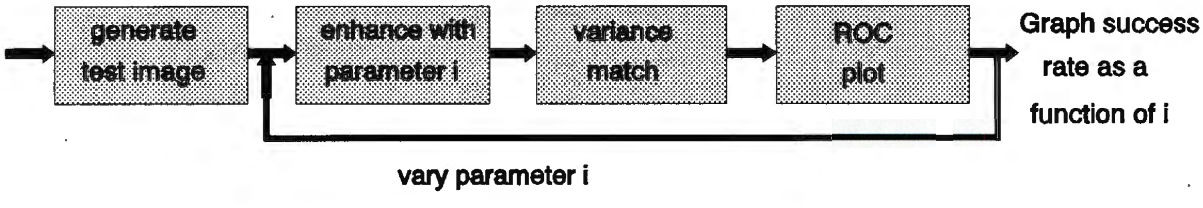


Figure 6.24: Block diagram depiction of an experiment for determining the most suitable enhancement technique in an automated nodule detection system.

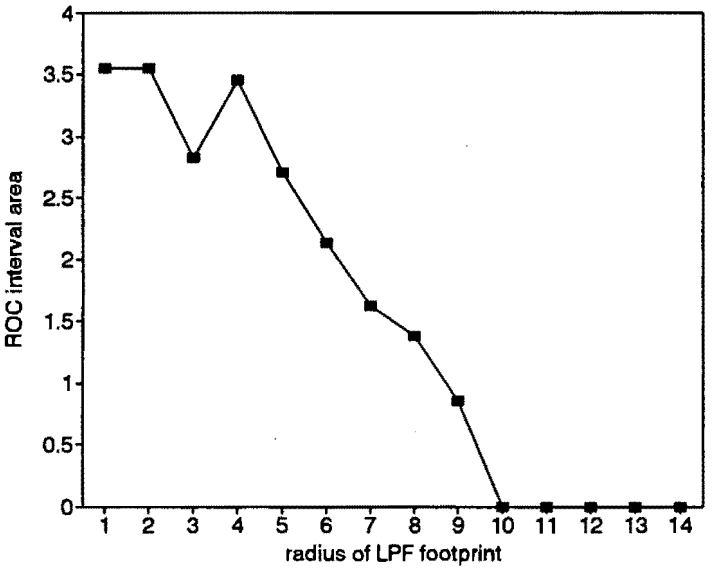


Figure 6.25: Graph of ROC interval area versus the radius of the linear filter footprint used in unsharp masking.

The footprint dimensions, for which the graphs are a maximum, are understood by examining the areas of the footprints in comparison with the nodule area. The area of a median filter footprint must be just greater than twice the area of the object which is to be enhanced. The pixel area of a nodule with radius of 8, is $\pi(8.5)^2 = 227$. The table below contains the pixel areas of a circular and a square footprint for a range of footprint radii.

footprint radius	square area	disc area
9	361	283
10	441	346
11	529	415
12	625	490

Twice the nodule area is 454 and therefore the median filter radius which is expected to give the peak response is marginally greater than 10 for the square footprint and between 11 and 12 for the circular footprint. These expected values correlate very well with the coordinates of the peaks read from the graph.

If the footprint dimension is larger than required, the enhanced image will contain more image data which has the effect of shifting the ROC plot to the right and thereby increasing the false hit rate.

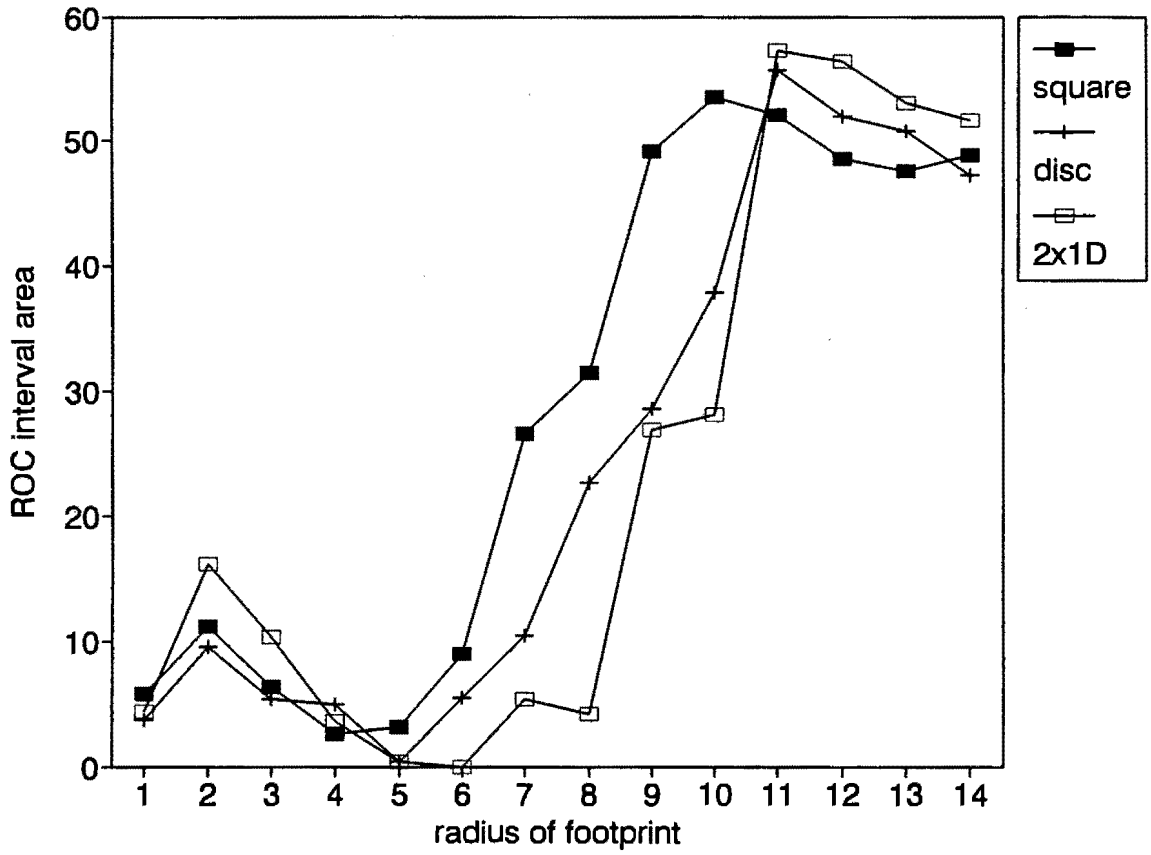


Figure 6.26: Graph of ROC interval area versus the radius of the median filter footprint used in unsharp masking.

The minor peak at a footprint radius of 2 might be due to the edge detecting effect of a small footprint.

6.5.2 Suitability of morphological structuring elements to automatic nodule detection

A top-hat transform can be likened to a filtering operation where the structuring element (SE) is analogous to the filter. The SE is defined by two parameters, *viz.* shape and dimension. This section discusses the evaluation of different SEs in terms of both shape and size.

In this experiment the objective was to determine (a) which SE, and (b) what dimension of SE is the most suitable for use in the automatic detection of nodules.

The test image (see section 6.2) was processed using a top-hat transform (see section 3.5) with a variety of SEs. This process was repeated using a range of sizes of the SE. Five SE shapes were tested in this manner. The results of ROC interval area versus SE radius were graphed for each of the SEs (see Figure 6.27).

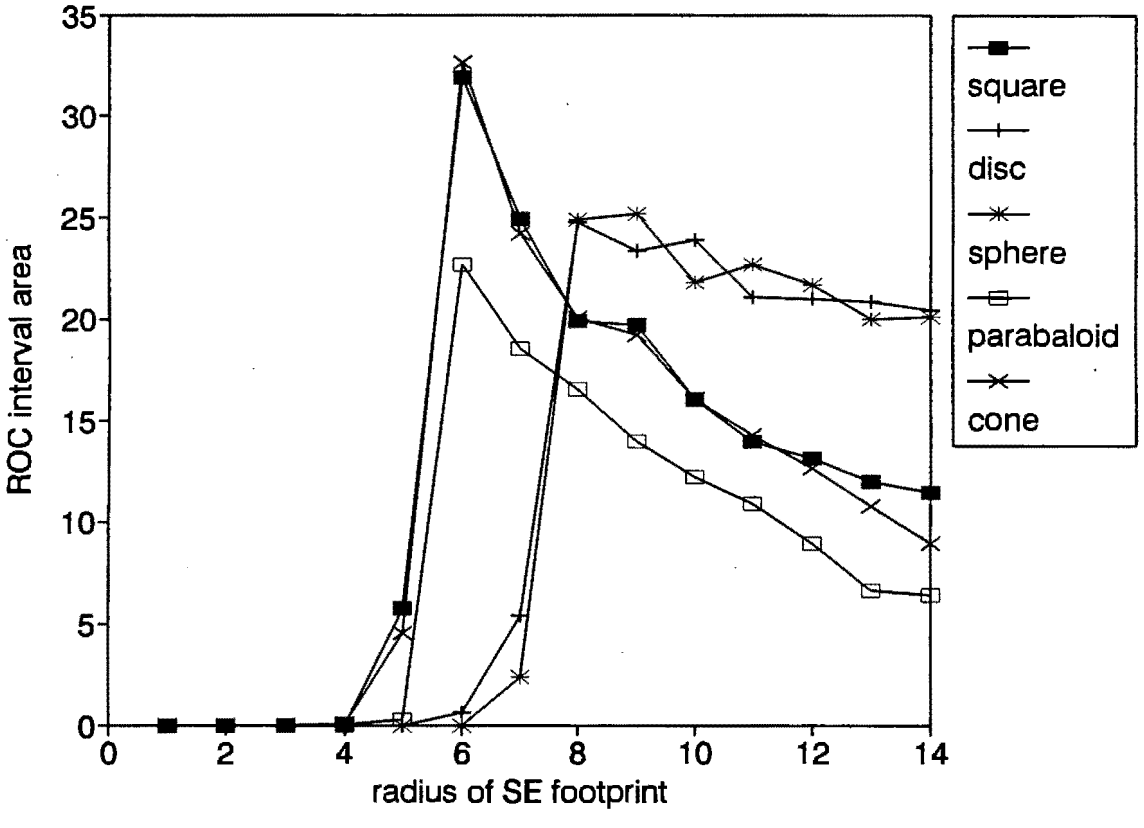


Figure 6.27: Graph of ROC interval area versus the radius of various structuring elements.

The graphs for all the SEs exhibited a trend which was anticipated. The graphs following this trend have three noticeable segments. Initially, when the SE radius is much smaller than the nodule radius, there is a very poor response and the ROC interval area is minimal. Then, as the radius of the SE approaches the radius of the nodule model, there is an abrupt increase in ROC interval area culminating in a peak. The final graph segment shows a slow decrease in ROC interval area as the SE radius becomes increasingly larger than the nodule radius. The first graph segment can be attributed to the almost total elimination of the nodule by a SE which *fits* into it (remembering that although the nodules in the image are largely eliminated, so too is the template). As the SE can no longer *fit* into the nodule, the nodule

is progressively better retained in the top-hat transformed image and subsequently matched well in the variance matching phase. Once the SE becomes large enough to retain the nodule, in the top-hat transformed image, the detection rate is maximised. An increase in SE radius has the effect of retaining larger structured noise, thus confusing the matching process. In the limit of a very large SE radius, the input image is not altered by the enhancement stage, and it is known that variance matching does not work effectively when the object to be detected is superimposed on large scale structures[5].

The SEs investigated were a flat disc, a flat square, a sphere[6], a cone, and a parabaloid. As mentioned above, all of these SEs exhibited the trend discussed. The responses of the SEs differed in the position of their peak. In particular, the square SE, cone SE and parabaloid SE graphs peaked when their radii were two less than the radius of the nodule model. The disc SE and sphere SE graphs peaked when the SE radius was just larger than the nodule radius. The premature peaking of the square, cone and parabaloid graphs can be attributed to the fact that all of these SEs have meaningful data at every point in the square footprint which they occupy. This is unlike the disc and sphere which only have meaningful data within the inscribed circle of their square footprints (see Figure 3.10). The disc and sphere can therefore *fit* into the nodule structure providing that it is less in radius. The other SEs (square, cone and parabaloid) will *not* fit into the nodule at an earlier stage[7].

From the graph it appears that the sphere or disc SE is the most suitable when the SE radius is equal, or larger, to the radius of the nodule sought.

6.5.3 The suitability of *peeling off* as an enhancement module

In section 3.4 the concept of *peeling off* was mentioned. In this section an experiment which determines the suitability of peeling off as a preprocessor is discussed. An image is said to have been peeled if it has been iteratively processed using the rolling ball algorithm (see section 3.5). In principle, the best simulation of peeling off is numerous iterations of a very small spherical SE. Unfortunately, small SEs lose their integrity as spheres due to the quantisation effect. Clearly, a compromise between SE size and number of iterations must be found. The purpose of this experiment is thus twofold. Firstly, this compromise must be found, and secondly, the suitability of peeling off as an enhancement technique must be evaluated. The

[5]Variance matching does not satisfy additive invariance (see section 4.4.4).

[6]Often referred to as the *rolling ball*.

[7]The diagonal of a square footprint exceeds the nodule diameter when the radius of the footprint is 6. A footprint of radius 6 has dimension 13x13 and consequently a diameter of 18. The diameter of nodule, radius 8, is 17.

results are shown in Figure 6.28.

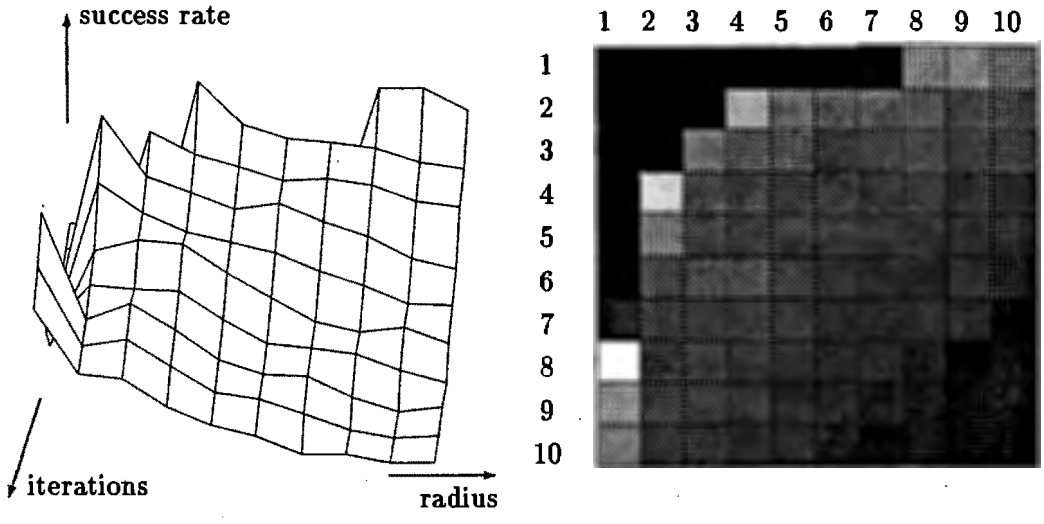


Figure 6.28: Surface plot and intensity map of the ROC interval area as a function of ball radius (horizontal variable) and number of iterations (vertical variable). The degree of success is proportional to the *whiteness* of the pixels in the intensity map. A maximum success measure of 32 occured when the radius was one and the number of iterations was eight.

From the results of this experiment, it is observed that multiple iterations with a small sphere produces a higher success measure than few iterations with a larger radius sphere. A hyperbolic trend is noticeable in the intensity map. This trend results because the success measure is approximately proportional to the product of the radius and the number of iterations.

6.6 Evaluation of the most suitable Analysis Module

Chapter 4 investigated template matching techniques, using correlation measures and the Hough transform circle finder, as the detection stage in the A.N.D system. This section attempts to evaluate the image analysis modules and to determine which is the most suitable.

In the previous section the enhancement techniques were evaluated as preprocessors to variance matching. Based on the results of that section three enhancement implementations were selected as the most successful. They are, (i) median unsharp masking with a circular footprint of radius 11, (ii) top-hat transformation with a disc SE of radius 8, and (iii) peeling off using eight iterations with a sphere of radius 1. As mentioned previously, exhaustive testing of all the combinations of enhancement and analysis modules is impractical. Thus, in order to gain some insight into the operation of the different analysis modules, three of the analysis

procedures were chosen to be used as the detection stage for the images enhanced by the aforementioned methods. ROC plots for the enhanced image followed by normalized correlation, the greyscale Hough circle finder and variance matching are superimposed on a single set of axes for comparison purposes. This is done for each of the three enhanced images and the ROC plots are shown in Figures 6.29, 6.30 and 6.31.

The labels in the graph legend **nc**, **gh**, **vm** and **50%** correspond to normalized correlation, greyscale Hough circle finder, variance matching, and the exponential curve (see section 5.3.1), respectively.

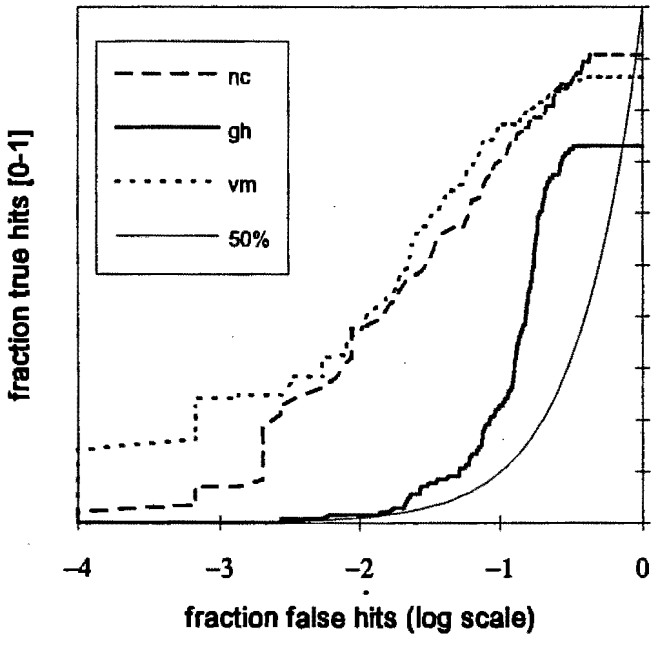


Figure 6.29: ROC plots of the different analysis techniques applied to the median unsharp masked image.

It is observed from the plots that no single analysis module produces the best results for all the preprocessed images. An optimal implementation of template matching is not a trivial exercise. A particular combination of shape and scale information must be determined if the process of template matching is to yield the maximum detection rate.

The highest success measure, 57, occured when median unsharp masking preceded variance matching. On that ROC plot, Figure 6.29, the percentage of true hits was 77% when the false hit rate was 10%. The total number of sites in the image was 6319[8], giving the total number of false sites as 6319 – 141 = 6178 (there are 141 true sites). Therefore 10% of the

[8]Derived from the total site formula $\lceil \frac{600}{8.5} \rceil . \lceil \frac{750}{8.5} \rceil = 6319$ (see section 5.3.1).

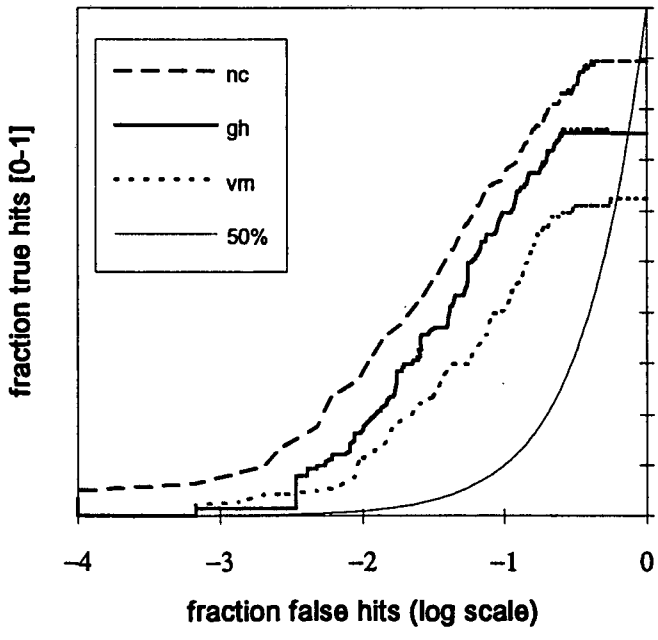


Figure 6.30: ROC plots of the different analysis techniques applied to the top-hat transformed image.

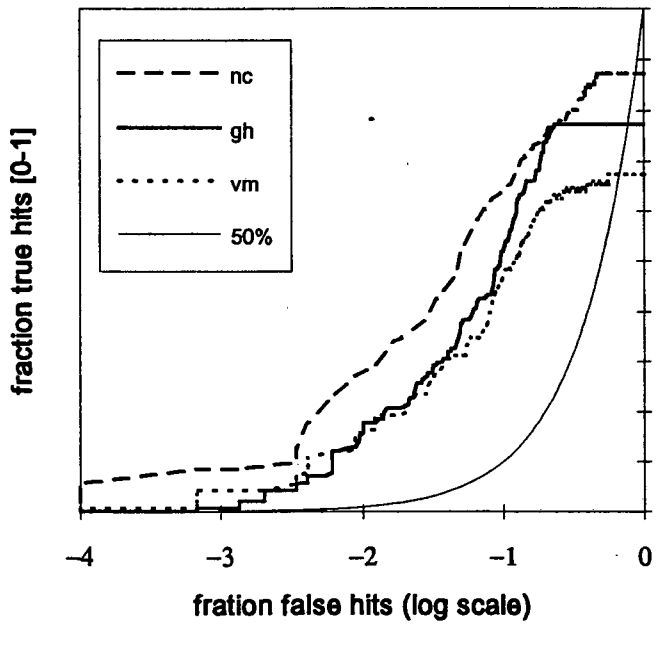


Figure 6.31: ROC plots of the different analysis techniques applied to the image which has undergone the peeling off process.

false sites is 618, and 77% of the true sites is 109. This indicates that only 1 out of every 7 sites, isolated by the computer as suspect, are true sites. For a false hit rate of 1%, the true hit rate was 37%. Therefore there are 62 false sites and 52 true sites in the image. The statistics improve when the false hit rate is 0.1%. At this rate the percentage of true hits is 24. Thus there are 34 true sites and only 6 false sites. From these statistics it is clear that the best combination of A.N.D modules results in a poor ratio of true hits to false hits when the number of false sites is much greater than the number of true sites.

For the very low false hit rate of 0.1% the method of variance matching is the only analysis technique which achieved a true hit rate greater than 10%, albeit in only one of the three examples. Time limitations prevented a more comprehensive examination of the most suitable analysis technique, but it does appear that the suitability of the analysis processes is also a function of the tolerable false hit rate.

Chapter 7

Conclusions and Recommendations

It was stated in an introductory paragraph of chapter 5, that "Humans are considerably more adept at object recognition, and/or classification, than are computers. However, even humans find the task of classifying nodules in chest x-rays particularly difficult, which presupposes that the ability of computers to perform proficiently is minimal." The possibility of assisted detection by the computer is further diminished because the image with which the computer works is a degraded version of the medical x-ray film which the diagnostician inspects. The degradation occurs in the scanning process by the digitisation which quantises both the spatial resolution and the intensity range.

Humans can only discern a maximum of 64 greyscale intensities at any one time, and therefore dynamic range reduction is necessary for human interpretation. The computer, on the other hand, has an advantage over humans in that a greater dynamic range improves the computer's understanding of the image.

In my opinion, dynamic range reduction for improved human perception is the only way the computer can assist the human in visually interpreting the image. The techniques for achieving human visual enhancement using dynamic range reduction often accomplish the necessary preprocessing which is required if the processed image is to be used as the input to an automated detection algorithm. In other words, enhancement for visual understanding and enhancement as a preprocessor have very similar requirements. Enhancement for subsequent automated detection does not require dynamic range reduction, but this is often a result of the enhancement process.

There are, in general, numerous parameters for each enhancement technique. It is expected that the optimal implementation of each of the enhancement techniques will not have been

achieved. Also, only educated guesses can be made when optimising nonlinear processes, such as median unsharp masking and the top-hat transform, as these operations are not well understood. Linear unsharp masking and the design of the optimal linear filter are, however, well understood procedures and it is therefore recommended that the optimal linear filter, for this application, be designed because it can act as a benchmark against which the nonlinear techniques can be judged. An exhaustive search of the multi-dimensional parameter space would be needed if the optimal implementation of each human visual enhancement technique were to be found. Even with exhaustive testing, the decision of the optimal technique is subjective, rendering it difficult, if not impossible, to evaluate. To further complicate matters the techniques are picture dependent. The enhancement for preprocessing has the advantage that it can be quantifiably assessed in combination with detection and classification modules. The disadvantage, though, is the massive increase in search space if exhaustive testing of the enhancement, analysis and classification is made.

The field of greyscale morphology shows promise and deserves more thorough investigation. The rotating rod algorithm, essentially a morphological process, was originally designed with the fulcrum of the rod in the centre. This design was later discarded because of noticeable artifact production. The replacement design, with the fulcrum at the end, does not generate obvious artifacts, but it has the severe limitation of being unable to enhance nodules on slopes, and it is consequently not additively invariant. In retrospect it might have been worth investigating the original design of the rotating rod, which is able to enhance nodules on slopes. It is known that most of the investigated techniques (linear and nonlinear unsharp masking, statistical difference filtering, and adaptive histogram equalisation) also produce noticeable artifacts. Artifacts are not a cause for concern if the process is intended as a preprocessor to a detection stage and, more importantly, if the artifacts do not "confuse" the detection process.

Peeling off is an example of a technique which was limited by the discrete nature of the image. Improved performance is expected from all processes if the resolution of the image is finer, in other words if there are more pixels per nodule. This allows the nodule, the filters and the structuring elements to be better described.

Image warping, or registration, is the only deterministic technique which has the potential of generating the perfect difference image because it has the fundamental ability to eliminate all anatomical information. This class of processes warrants thorough investigation.

Edge detection is an important step in many image processing applications. Satisfying Canny's requirements and designing the optimal edge detector for a specific application is often a difficult task. The nonspread edge detector appears to do well but suffers the flaw of not having rotational invariance.

Template matching is used extensively in signal and image processing applications, but it is not the only means of detection. There are numerous variations on the template matching theme which are concerned with different descriptions of the object to be detected. For example, the Hough circle finder uses the edge information of the nodule, while normalized correlation uses the full three dimensional shape description of the nodule. Two important considerations of additive invariance and intensity matching are vital for optimal detection. Variance matching was an attempt at an implementation of template matching which allowed intensity matching. But variance matching was not additively invariant and consequently required the image to be preprocessed. Because the planar shape of an object in an image is its most recognisable feature, the Hough transform circle finder performs reasonably well on an edge detected image. It should be possible to increase the dimensionality of the Hough accumulator array, and in this way match using other significant characteristics of a nodule. An experiment which evaluated different template matching procedures using edge, shape and intensity information was non-conclusive in determining which of these descriptions is the most appropriate. A feature vector with weighting for features which are thought to characterize a nodule could be investigated. This suggests the use of a learning paradigm, such as self organising feature maps (a class of neural networks), for determining the most relevant features and their respective weighting.

The binary Hough transform circle finder only works as well as the greyscale Hough transform circle finder if the correct threshold value is chosen. A suitable threshold value is often very difficult to determine and therefore the greyscale circle finder is a more effective technique for use with real images.

The classification techniques employed were simplistic. Further discriminating measures could be introduced to lessen the false positive hit rate. Contextual information, for example the locations of the major arteries (which appear similar to nodules), could be included for discrimination. Circularity in the correlation surface could also be employed to decrease the likelihood of false hits. The use of multiscale measures is another area which warrants further investigation. The abrupt change and the distinctive shape in the graph of success rate versus footprint size make the use of multiscale measurements for discrimination an attractive option.

Determining a success measure from an ROC plot is an interesting problem of reducing dimensionality. It is likely that the ROC interval area used as the success measure is not the most appropriate means of quantifying the ROC plot for this application. Investigation into alternative methods of quantifiying the ROC plot should be made in conjunction with a radiologist.

The experiments conducted in chapter 6 indicate that the automated nodule detection (A.N.D) system *is* capable of detecting camouflaged nodules in digitised chest x-rays. It was, however,

also evident that even the best combination of A.N.D modules results in a poor ratio of true hits to false hits when the number of false sites is much greater than the number of true sites. Increasing the ratio of true hits to false hits can be achieved by improving the operation of the enhancement or analysis module, or by increasing the sophistication of the discriminating mechanisms in the classification process.

The recent trend towards the utilisation of digitised images in the medical environment suggests that implementations of enhancement and automatic detection algorithms are becoming indispensable. Although this thesis has not solved the problem of reliable automated detection of lung cancer nodules in chest x-rays, it has laid the foundations of a modular A.N.D system.

Bibliography

[1] K.M. Andress and D.L. Wilson. Optimization of Morphological Structuring Elements for Angiogram Enhancement. In Murray H. Loew, editor, *SPIE Medical Imaging V: Image Processing*, volume 1445, pages 6–10, San Jose, California, 1991.

[2] D. H. Ballard. Generalizing the Hough Transform to Detect Arbitary Shapes. *Pattern Recognition*, 13(2):111–122, 1981.

[3] Wayne Borchardt. Undergraduate thesis: Medical X-ray Enhancement: A computer-aided nodule detection algorithm for lung cancer diagnosis, November 1990.

[4] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, November 1986.

[5] Bruce Curtin. Undergraduate thesis: Subjective Testing to Determine the Average Number of Grey Levels that can be Distinguished on a Monitor by the Human Visual System, November 1991.

[6] Gerhard de Jager. Multidimensional Digital Signal Processing course notes, 1991.

[7] Department of Electrical and Computer Engineering, University of New Mexico, Alburquerque, NM. *KHOROS MANUAL*, 1991. Release 1.0, patch level 5.

[8] Richard O. Duda and Peter E. Hart. *Pattern Classification and Scene Analysis.* John Wiley & Sons, Inc, 1973.

[9] R.O. Duda and P.E. Hart. Use of the Hough Transform to detect lines and curves in pictures. *Communications of the ACM*, 1:11–15, January 1975.

[10] Michael P. Ekstrom, editor. *Digital Image Processing Techniques.* Academic Press, inc., 1984.

[11] J.P. Ericksen, S.M. Pizer, and J.D. Austin. MAHEM: A Multiprocessor Engine for Fast Contrast-Limited Adaptive Histogram Equalization. In *SPIE Medical Imaging IV: Image Processing*, volume 1233, pages 322–333, San Jose, California, 1990.

[12] M.L. Giger, K. Doi, and H. MacMahon. Image feature analysis and computer-aided diagnosis in digital radiography. 3.Automated detection of nodules in peripheral lung fields. *Medical Physics*, 15(2):158–166, 1988.

[13] Ernest L. Hall, Richard P. Kruger, and Samuel J. Dwyer. A Survey of Preprocessing and Feature Extraction Techniques for Radiographic Images. *IEEE Transactions on Computers*, c-20(9):1032–1044, 1971.

[14] H. J. A. M. Heijmans and C. Ronse. The Algebraic Basis of Mathematical Morphology: I. Dilations and Erosions. *Computer Vision, Graphics, and Image Processing*, 50:245–295, 1990.

[15] Bertil Jacobson and John G. Webster. *Medicine and Clinical Engineering.* Prentice-Hall, Inc, 1977.

[16] C. Kimme, D. Ballard, and J. Sklansky. Finding Circles by an Array of Accumulators. *Communications of the ACM*, 18(2):120–122, February 1975.

[17] Jae S. Lim. *TWO-DIMENSIONAL SIGNAL AND IMAGE PROCESSING.* Prentice-Hall, International, Inc., 1990.

[18] Petros Maragos and Ronald W. Schafer. Morphological Systems for Multidimensional Signal Processing. *Proceedings of the IEEE*, 78(4):690–710, 1990.

[19] K. Morishita and T. Akimbo. Image Registration Method Using Adaptive Nonlinear Filter. *Systems and Computers in Japan*, 19(9):41–50, 1988.

[20] Wayne Niblack. *An Introduction to DIGITAL IMAGE PROCESSING.* Prentice-Hall International (UK) Ltd, 1986.

[21] Lawrence O'Gorman and Arthur C. Sanderson. The Converging Squares Algorithm: An Efficient Method for Locating Peaks in Multidimensions. *IEEE transactions on pattern analysis and machine intelligence*, 6(3):280–288, 1984.

[22] Bernard Ostle and Richard W. Mensing. *Statistics in Research.* The Iowa State University Press, third edition, 1975.

[23] Kelly Rehm and William J. Dallas. Artifact Suppression in Digital Chest Radiographs Enhanced with Adaptive Histogram Equalization. In *SPIE Medical Imaging III: Image Processing*, volume 1092, pages 290–300, 1989.

[24] Azriel Rosenfeld and Avinash C. Kak. *Digital Picture Processing*, volume 2. Academic Press, inc., second edition, 1982.

[25] Jean Serra. Introduction to Mathematical Morphology. *Computer Vision, Graphics, and Image Processing*, 35:283–305, 1986.

[26] J. Song and E. J. Delp. The Analysis of Morphological Filters with Multiple Structuring Elements. *Computer Vision, Graphics, and Image Processing*, 50:308–328, 1990.

[27] Stanley R. Sternberg. Grayscale Morphology. *Computer Vision, Graphics, and Image Processing*, 35:333–355, 1986.

[28] Cathy Stoner, April 1992. Specialist Radiologist, New Groote Schuur Hospital, Private communication.

[29] Ferrel G. Stremler. *Introduction to Communication Systems.* Addison-Wesley Publishing Company, second edition, 1982.

[30] Paul Symonds. The investigation of the characterisation of flotation froths and design of a machine vision system for monitoring the operation of a flotation cell ore concentrator. Master's thesis, University of Cape Town, 1992. not yet completed.

[31] Harry Wechsler. *Computational Vision.* Academic Press, Inc., 1990.

Appendix A

Source Code Description

The relevant source code to this thesis is provided on a single floppy disk[1]. All code was written in C on the SUN SPARCstation 2.

There are 3 models of the program that can be built. Each model deals with images of different dimensions. The BIG model currently uses images of 720 rows by 1800 columns and 12 bits per pixel. The MED model operates on images of 600 by 750 by 12 bits. The SMALL model operates on the images provided by most framegrabbers, *viz.* 512 by 512 by 8 bits.

The building of the programs from the source code is facilitated with UNIX makefiles. There are 3 such makefiles: **BIG.mak, MED.mak, SMALL.mak.** The built programs will have the names **b_demo, m_demo** and **s_demo.**

The source code is logically divided into separate files. The highest level file **x_demo.c** provides a simple user interface from which all of the processes can be executed. The source code files all have associated header files which contain a list of the functions which are publically available[2] after linking. The source code files are:

- **x_curve.c** and **x_curve.h** for nonlinear adaptive curve fitting.
- **x_edge.c** and **x_edge.h** for all edge detection routines.

[1] A large amount of code was written because image processing packages, such as **Khoros** [7] and **Sunvision** [30], were not available at the time of implementation. Despite the quantity of code, this thesis was not a software project and therefore it was deemed unnecessary to include source code listings.

[2] This code was not written in C++ and therefore public, or private, functions are only a suggestion by the author.

- **x_eqlise.c** and **x_eqlise.h** for all histogram modification routines.
- **x_fft.c** and **x_fft.h** for Fast Fourier Transform routines.
- **x_filter.c** and **x_filter.h** for filtering functions, linear, nonlinear and adptive.
- **x_hough.c** and **x_hough.h** for the Hough transform circle finder routines.
- **x_io.c** and **x_io.h** for all file input/output routines.
- **x_match.c** and **x_match.h** for all the template matching processes, excluding the Hough transform.
- **x_morph.c** and **x_morph.h** for all the standard morphological processes.
- **x_rod.c** and **x_rod.h** for all implementations of the rotating rod algorithm.
- **x_sort.c** and **x_sort.h** for classification related functions, including the generation of the ROC plot.
- **x_utils.c** and **x_utils.h** for all general image processing utilities.
- **stats.c** and **stats.h** for statistical functions.

All of the built programs include the additional header **x_gen.h** which provides important definitions and conventions that are used throughout the code. In addition to this header, the program **b_demo** all includes **b_.h** and **b_types.c**. **b_.h** defines the number of rows and columns and the number of intensity levels of the images on which **b_demo** operates. **b_types.c** defines the data types which will be used throughout the code. This includes the defintion of the data type for an image, amongst others. Similar files **m_.h**, **m_types.c** and **s_.h**, **s_types.c** exist for the programs **m_demo** and **s_demo**, respectively.

The program may be rebuilt to operate on any image of any dimensions by editing the **b_.h** (or **m_.h** or **s_.h**) file and recompiling with the use of the makefile.

SunSoft INTERACTIVE UNIX

System V/386 Release 3.2, Version 3.0

PRODUCT SPEC

SunSoft's INTERACTIVE UNIX System V/386 Release 3.2 Operating System, Version 3.0 provides users with a powerful and flexible working environment. A friendly user interface and tutorial-style documentation make the system easy to learn.

The SunSoft INTERACTIVE UNIX operating system is the foundation of SunSoft's current 80x86 Product Family and the path to Solaris 2.0 on Intel. It is a multi-user, multitasking system based on Release 3.2 of UNIX System V. It conforms to the accepted standard for UNIX implementations, the System V Interface Definition of UNIX System Laboratories, Inc. (USL), a subsidiary of AT&T and is fully compatible with other USL-certified UNIX System offerings.

The SunSoft INTERACTIVE UNIX Operating System includes a number of enhancements to UNIX System V Release 3.2.

Performance Improvements

INTERACTIVE UNIX includes two features that improve data throughput:

- **High Performance Device Driver (HPDD).** The HPDD provides improved disk throughput and allows several drivers to share a hardware controller.
- **Fast File System (FFS).** The FSS is an enhancement to the UNIX file system support code in the System V kernel that improves disk file I/O performance.

Internationalization Support

Standards conformance continues to be a major areas of focus for SunSoft. With Version 3.0, strict conformance to the current version of the X/Open Portability Guide (XPG3) has been added to the existing compliances with POSIX 1003.1 and SVID Issue 2. An International Supplement is available to support localized collation sequences, internationalized message files, localized regular expression handling and localized data and monetary formats.

It also includes over 30 internationalized commands.

Easy to Install, Administer, and Use

Easy-to-follow procedures guide you through every step of the installation process with pull-down menus, pop-up windows and forms — in color. Context-sensitive on-line help is available at all stages of system administration, with the result that most day-to-day administration can be performed without a manual.

Supports a Wide Variety of Hardware

SunSoft INTERACTIVE UNIX supports most popular 80386/80486 ISA, EISA, and MCA hardware platforms. Refer to the *SunSoft INTERACTIVE UNIX Hardware Compatibility Guide* for a complete list of supported devices.

SunSoft
A Sun Microsystems, Inc. Business

Version 3.0 enhances EISA bus systems support by providing:

- Support for physical memory over 16MB, up to a maximum of 256MB
- Support for a number of EISA disk controllers and SCSI adapters
- A kernel driver for developers of EISA controllers to interface to the EISA NVRAM.

It also supports a variety of peripheral controllers including: multiport boards, intelligent and non-intelligent network boards, video display adapters, disk controllers, 9-track and cartridge tape controllers, and pointing devices.

Product Highlights

All the Components of UNIX System V Release 3.2

- Base UNIX system
- Kernel extension
- Basic utilities
- Advanced utilities
- System administration
- Terminal interface extension
- Virtual terminal feature
- STREAMS
- XENIX compatibility

Easy Installation and Administration

- Menu-driven installation and menu-driven System Administration
  - CUI-based
  - Pulldown menus, pop-up windows, screen forms
  - Context-sensitive help
  - Error logging
  - Keystroke logging
  - Consistent operations
- Menu-driven Kernel Configuration utility
- Modular documentation

Standards Compliance

- True UNIX System V Release 3.2
- SVID (passed SVVS)
- FIPS 151-1
- POSIX 1003.1 (passes NIST PCTS)
  - Supports job control and job control shell
  - Job control support added to utilities
  - Data interchange utilities (pax)
  - Supplementary documentation
- C-2 Security Extension (optional)
- *X/Open Portability Guide, Edition 3*

Increased Performance

- Disk I/O up to 300% faster than standard UNIX
- Very Fast File System
- High Performance Disk Driver
- Maintains System V file system format
- ST506™, RLL, ESDI and SCSI drives supported

DOS Integration

- Retain DOS and UNIX files in separate hard disk partitions
- Move files between DOS and UNIX
- Access DOS files from UNIX applications
- Integrate DOS and UNIX files into single file system with optional extensions

International Supplement

- 30 utilities modified to support operations in international environment
- Keyboard description files for European keyboards
- Collation tables tools
- Utilities for message catalogs
- International Supplement Guide

Supports Wide Variety of Hardware

- 80386/80486-based ISA, EISA, or MCA bus
- Enhanced EISA system support
- Multiport boards
- Intelligent and non-intelligent network boards
- Video display adapters
- Disk controllers
- 9-track and cartridge tape controllers
- Pointing devices
- CD-ROM/WORM

Berkeley Features

- Berkeley *sendmail* Facility — an internetwork mail interface configured for UUCP, TCP/IP and other networks
- Berkeley C Shell

Documentation

- Modular documentation structured to meet the needs of new and experienced UNIX users
- Guide for New Users
- On-line manual pages

Minimum Hardware Requirements

- 80386 or 80486 IBM-compatible microcomputer, with ISA, EISA, or MCA bus
- 4 MB RAM
- 40 MB hard disk space
- 3 1/2" or 5 1/4" floppy drive
- Video adapter card (MDA, Hercules™, EGA, VGA or high-res)
- Monitor
- 101- or 102-key enhanced AT keyboard

SunSoft, Inc. • 6601 Center Drive West, Suite 700 • Los Angeles, CA 90045 • (800) 346-7111 • (310) 348-8649 • FAX: (310) 348-8600
Federal Systems • SunSoft, Inc. • 2100 Reston Parkway • Suite 101 • Reston, VA 22091 • (703) 758-0900 • FAX: (703) 620-1244
European Headquarters • SunSoft, Inc. • St. John's Court • Easton Street, High Wycombe • Buckinghamshire HP11 1JX • United Kingdom • (44) 494-472900 • FAX: (44) 494-472383

SunSoft
A Sun Microsystems, Inc Business

Printed in USA 1/92 915125-01