IMPROVING THE PRECISION OF LEG ULCER AREA MEASUREMENT WITH ACTIVE CONTOUR MODELS

Timothy David Jones

A submission presented in partial fulfilment of the requirements of the University of Glamorgan/Prifysgol Morgannwg for the degree of Doctor of Philosophy

This work was carried out in collaboration with the Royal National Hospital for Rheumatic Diseases, Bath and the University Hospital of Wales, Cardiff

May 1999
“But a Samaritan, as he travelled, came where the man was; and when he saw him, he took pity on him. He went to him and bandaged his wounds, pouring on oil and wine.”

LUKE 10:35
ACKNOWLEDGEMENTS

I extend warm and hearty gratitude to Prof. B. F. Jones for his supervision and helpful guidance throughout the duration of this project. His careful oversight and concern have been of great benefit to me through a long and sometimes arduous journey. I deeply appreciate the patience and understanding he has shown and in particular for inspiring me to focus the work and complete this thesis.

I would also like to take this opportunity to extend my thanks to both of my second supervisors, Dr. Peter Plassmann and Dr. S. J. Rees, for providing helpful suggestions and advice.

I gratefully acknowledge the assistance of Prof. E. F. J. Ring of the Royal National Hospital for Rheumatic diseases in Bath and Dr. K. G. Harding of the University Hospital of Wales, Cardiff for allowing access to patients and for providing images of leg ulcers.

I am also grateful to Dr. Wyn Williams for his help with variational calculus and to Dr. David Jones for assisting in the design of the manual delineation experiments. Thanks also to Dr. Jamal Ameen for casting a critical eye over the statistical analyses.

This work was partly funded by the Higher Education Funding Council of Wales (HEFCW) under the Development of Research (DevR) initiative.
A leg ulcer is a chronic wound of the skin that, at best, takes many months to fully heal and causes great distress to the patient. Treating leg ulcers places a large financial burden upon the National Health Service in the United Kingdom, estimated to be in excess of £300M annually. Measurement of the size of leg ulcers is a guide to assessing the progress of wound healing, and the use of non-invasive measurement techniques avoids damaging or infecting the wound. The area of a leg ulcer is currently measured by presenting a human observer with a captured video image of a wound, who then uses a mouse or pointing device to delineate the wounded region. Typically, the standard deviation of area measurements taken this way is approximately 5% of the wound area. In addition, different observers can show a bias difference in their area measurements from 3% to 25% of the wound area. It is proposed to reduce the incidence of such errors by using an active contour model to improve the delineation. Four different models are developed by adapting and applying several contributions made to the active contour model paradigm. Novel features include an external force that acts normally, but not tangentially, to the boundary, a new external energy term that promotes homogeneity of the gray level at the edge of the wound and the application of the minimax principle for setting the parameters of an active contour model with piecewise b-spline curves. The algorithms provide the physician with a new and practical tool for producing area measurements with improved precision and are semi-automatic, requiring only a manual delineation to start the algorithm. In most cases, measurement precision is improved by application of the algorithms. Many wounds give rise to measurable bias differences between average manual area measurements and the corresponding algorithmic area measurements, typically averaging 3% to 4% of wound area. With some wounds the bias magnitude can exceed 10% as a result of the contour partly deviating from the true edge of the wound and following a false edge.
CONTENTS

1 Introduction

1.1 Leg Ulcers: Causes and Concerns 1
1.2 Background to the Thesis 3
1.3 Practical Wound Measurement 4
1.4 Proposed Solution 5
1.5 Contribution of this Work 7
1.6 Definition of Measurement Terms 8

2 Literature Survey

2.1 Image Segmentation Methods 11
   2.1.1 Histogram Analysis 12
   2.1.2 Region Growing and Split-and-Merge Schemes 14
   2.1.3 Colour-Space Analysis and Segmentation 15
   2.1.4 Pixel Classification for Segmentation 16
   2.1.5 Gradient Operators, Edge and Zero-Crossing Detectors 17
   2.1.6 Watershed Segmentation 20
   2.1.7 Radial Search Algorithms 21
   2.1.8 Boundary Following Algorithms 21
   2.1.9 Summary 22

2.2 Active Contour Models 23
   2.2.1 Modelling of the Contour 23
   2.2.2 Formulation of the Energy Functional 24
2.2.3 Minimisation of the Energy Functional 25
   2.2.3.1 Elastic Energy and Forces 25
   2.2.3.2 Bending Energy and Forces 27
2.2.4 Weaknesses of the Kass Model 28
2.2.5 Solution of Euler-Lagrange Equations 29
2.2.6 Spatial Discretisation 30
   2.2.6.1 Finite Element Shape Functions 31
   2.2.6.2 Stiffness Matrix Equations 34
2.2.7 Time Discretisation 34
2.2.8 Stability of the Iterative Solution Method 36
2.2.9 Direct Energy Minimisation Approaches 37
2.2.10 Element Collapse Problem 38
2.2.11 Modified Internal Energy and Forces 39
2.2.12 Region Pressure Energy 41
2.2.13 Directional External Energy and Forces 41
2.2.14 Setting the Contour Regularisation Parameters 43
2.2.15 Initialisation of Active Contour Models 45
2.2.16 Global Energy Minimum Seeking 46

2.3 Summary 48

3 Algorithm Development

3.1 Producing the External Potential Field 50
   3.1.1 Depletion of the External Energy and Attractive Forces 51
   3.1.2 Translation of Edges 53
   3.1.3 Scale Descent Algorithm 55
3.2 The Ground-Attract (GA) Algorithm 56
   3.2.1 Stability of the Solution 59
   3.2.2 Properties of the Model 59
   3.2.3 Minimisation Algorithm 61
   3.2.4 Possible Improvement of the Model 62
3.3 The Tangent-Normal (TN) Algorithm
3.3.1 External Force Made Normal to the Contour
3.3.2 Elastic Force Made Tangential to the Contour
3.3.3 Implementation of Tangential-Only Elastic Force
3.3.4 Stability of the Solution
3.3.5 Minimisation Algorithm
3.4 Normalisation of Parameters
3.5 The Minimax Algorithms (MX and MG)
3.5.1 Continuity Term
3.5.2 Curvature Term
3.5.3 Gradient External Energy Term
3.5.4 Gray Level External Energy Term
3.5.5 Continuous Domain Minimisation
3.5.6 Iteration and Convergence of Direct Minimisation Algorithms
3.6 Summary

4 Manual Delineation Experiments

4.1 General Assessment of Manual Delineation Performance
4.1.1 Objectives
4.1.2 Experimental Set-up and Procedure
4.1.3 Hypothesis Test for Bias: ANOVA Fixed Effect Model
4.1.3.1 The ANOVA Bias Test Model
4.1.3.2 Pre-ANOVA Checks for Violation of Assumptions
4.1.3.3 Inspection of Images for Regions of Ambiguity: Median Boundary Method
4.1.4 Model for Delineator Precision
4.1.5 Results

4.1.5.1 Summary of Pre-ANOVA Check Results 94
4.1.5.2 Variance Analysis 95
4.1.5.3 Bias Analysis 95
4.1.5.4 Test for Overall Delineator Bias 98
4.1.5.5 Precision Estimates 101

4.1.6 Discussion 102

4.1.7 Conclusions 104

4.2 The Effect of Image Display Size on Delineation Accuracy 106

4.2.1 Objectives 106
4.2.2 Development of Regression Model for Bias 106
4.2.3 Development of Regression Model for Variance 113
4.2.4 Experimental Set-up and Procedure 115
4.2.4.1 Creation of Phantom Wound Image 116
4.2.5 Bias Model Results 116
4.2.6 Variance Model Results 119
4.2.7 Discussion 122
4.2.8 Conclusions 125

PLATES 1 & 2    Original Colour Plates of Wound Images
inserted between pages 125 and 126

5    Parameter Setting Experiments

5.1 Introduction and Preliminary Notes 126
5.1.1 A Note on Varying the Initial Contour Position 127
5.1.2 Calculation of Contour Area 128
5.1.3 Calculation of Area Precision 128

5.2 Contour Regularisation Parameters 129
5.2.1 Range and Resolution of Parameter Levels 130
5.2.2 GA Algorithm: Parameter Varying Results 131
6 Active Contour Model Experiments

6.1 Objectives 149
6.2 Experimental Procedure 151
   6.2.1 Additional Image Set 151
   6.2.2 Seeding the Algorithms: Manual Delineation 151
6.3 Development of Analytical Procedures 152
   6.3.1 A Note on the Application of Statistical Methods 152
   6.3.2 Improvement of Manual Results 153
   6.3.3 Development of Graphical Procedure 155
   6.3.4 Comparison of Algorithms: Overall Test for Precision 156
   6.3.5 Quantification of Bias and Precision 158
6.4 Experimental Results 161
   6.4.1 Detection and Analysis of Outliers 161
   6.4.2 Relative Improvement of Algorithms 163
   6.4.3 Overall Precision Test Results 166
   6.4.4 Bias and Precision Results 167
6.5 Summary 169
7 Discussion

7.1 General Performance of Algorithms
7.1.1 Precision Measurement 171
7.1.2 Bias Measurement 172
7.2 Analysis of Failure Cases 176
7.2.1 Cases of Both Failure Modes 176
7.2.2 Insignificant Cases of Variance Ratio 177
7.2.3 Cases Where Bias is Significant 178
7.2.4 Identifiable Causes of Bias Errors 179
7.3 Consequences of Errors 180

8 Conclusions

8.1 Manual Performance 183
8.2 Algorithm Performance 184
8.3 General Statement 185
8.4 Limitations of this Work 187
8.5 Recommendations for Further Work 187

Appendices

A Formulation of B-Spline and Hermite Finite Element Models A-1
B Manual Delineation Experimental Data A-7
C Active Contour Model Experimental Data A-8

References A-10
Chapter 1 Introduction

1 INTRODUCTION

1.1 Leg Ulcers: Causes and Concerns

The skin is a complex organ of the human body and its property of self-regeneration and its contribution to thermoregulation (temperature control of the body) are often taken for granted. The skin is a first line of defence against infection carried by myriad agents with which it is in daily contact. It prevents foreign bodies from penetrating the underlying tissues and organs and protects against damage caused by harsh physical contact. Wounds of the skin are caused by many different factors both internal and external to the human frame. Common injuries caused by external factors include cutting, abrasion, bruising and burning. Injuries that tear or remove portions of the skin are classified as acute wounds and the natural remedial action of healthy skin is to close the wound. In contrast, the occurrence of a leg ulcer is a symptom of an underlying problem with the skin and its blood supply and drainage systems. Leg ulcers may be triggered by external agents: even something normally considered harmless, such as a sharp blow against an unyielding object, can precipitate the onset of a leg ulcer. They may also develop without external stimuli due to the severity of the underlying condition of the skin. Since the major factor in leg ulcers is a degeneration of the dermis itself, they are classified as chronic wounds and are typified by a much extended healing time when compared to acute wounds.

All wounds that heal are healed by the skin's own in-built mechanisms, without which no healing could occur. In the case of acute wounds, healing can be rapidly aided by primary intention, generally the physical closure of the wound by sutures. Such action relies upon the mechanical strength of the adjacent skin tissue to resist being torn by the suture material. A leg ulcer may be considered to be a symptom of the breakdown of the skin's ability to adequately heal itself. The poor condition of the skin in the vicinity of
an ulcer means that closure of the wound by primary intention is not possible and, in any event, this would only be treatment of the symptom rather than treatment of the underlying cause of the ulcer. Therefore wound healing must be aided by the external influence of medical treatment and supervision. This practice is termed healing by secondary intention.

Healing of a wound is a relatively slow process despite the attentions of the best understanding, although it should be noted that without appropriate medical treatment a leg ulcer will almost certainly never heal. The prevalence of leg ulcers in the UK alone has been estimated to affect 1-2% of the total population and estimates of the total cost of treating leg ulcers have recently exceeded £300 million (Davis et al., 1992). A typical treatment regime may be based on a combination of the following interventions:

- Application of a dressing to absorb wound exudate and keep the wound moist to promote healing.
- Applying a pressure bandage to encourage movement of lymphatic fluids and blood through the veins.
- Cleansing of the wound to reduce the possibility of infection. (Microbiology has revealed that infecting agents, e.g. staphylococcus aureus, colonise the human skin and even leg ulcers without causing 'infection'. The infection occurs when the colony is allowed to expand to much greater proportions).
- Where necessary the patient may be surgically treated to increase vein throughput.

In time, a healing wound will tend to reduce in size, whereas the size of a deteriorating wound will increase. Wounds may take several months to heal fully, and even then they may recur because of persistence of the underlying conditions in the dermis which first caused the ulcer. Other wounds may exist for more than a year and sometimes never fully heal and it is not unusual for an ulcer to be deemed 'indolent', indicating a stagnation of the healing process. Therefore, since changes in wound size are an
indicator of the direction of the healing process, measurement of a wound's physical dimensions is an important tool in the management and monitoring of the healing process. To determine whether a wound is changing in size, the physician must compare recent wound measurements with the current measurement and needs to be satisfied that the measured differences are indeed a manifestation of wound healing (or conversely, deterioration) and not due to experimental error. Clearly, if there is no detectable difference, then either the measurement precision is poor leading to low confidence in measured difference or the wound is not responding to treatment and is thus indolent. The measurement precision therefore influences the likelihood of detecting changes in wound size. Generally, a patient’s wound will be inspected on a weekly or fortnightly basis when the existing dressing will be replaced, affording the opportunity to measure the wound with little interference. Failure to reliably detect a significant size change will lead to a delay of one or two weeks if one considers that the course of action may be altered partly as a consequence of measurement results.

1.2 Background to the Thesis

This project arose as an enhancement of the MAVIS instrument (Plassmann, 1992, Jones and Plassmann, 1995). MAVIS stands for Measurement of Area and Volume InStrument, and, as its name implies, this instrument has the capability of measuring the surface area of a wound and of estimating the wound volume defined as the volume of that space enclosed between the wounded surface and the reconstructed original healthy skin surface over the wound site. Measurement of both surface area and volume requires 3-D information in the form of a depth-map of the wound surface which the instrument obtains by projecting structured light on the wound and subsequently processing the resulting captured image. One step in this image processing sequence is the definition of the boundary of the wound which is required for measurement of the wound’s perimeter, area and volume. Currently, the boundary definition is performed by manually delineating the wound using a standard computer mouse to guide an on-screen cursor.
1.3 Practical Wound Measurement

Several methods for measuring various dimensions of a wound exist, ranging from an estimate of area based upon measurements of linear dimensions made with a modified ruler (Kundin, 1989) to video-capture and three-dimensional imaging techniques (Plassmann, 1995, Plassmann and Jones, 1998). Imaging methods make use of video capture which involves more expensive equipment but has the advantage of not making contact with the wound, i.e. they are 'non-invasive', thus alleviating the possibility of imparting infection or of injuring delicate tissue. Both of these events will prolong occurrence of the wound, perhaps by many weeks, so it is in the best interest of the patient that such events are avoided. This thesis is concerned with the application of computer-based imaging to the wound measurement problem. In this context, the most rudimentary procedure is to perform a manual tracing (delineation) of the boundary of a wound image by using either a track-ball or mouse. Once the boundary has been defined, the area and circumference can be measured in terms of pixel areas and lengths respectively. These measurements may then be converted to physical units by application of suitably derived pixel scaling factors, notwithstanding the many sources of error and distortion introduced by the imaging and video capture processes.

Manual delineation is subject to several criticisms that bear upon its expected performance as a measurement process and these are directly linked to the human factor. Due to the nature of hospital work, a successful measurement process needs to be fast in operation, at least in terms of capturing necessary data required for the measurement, and easy to use. Automation of the wound measurement process by replacing the task of manual delineation with a wound measurement algorithm would place less of a demand upon the physician. Removing the delineation process and replacing it with a computer-based algorithm removes the subjective element intrinsic to human judgment. Firstly, it is considered that the opinion as to the extent of a wound may well vary among manual delineators introducing the possibility of measurement bias. Secondly, the process of delineation is subject to the available level of human dexterity, which is likely to be variable, both for the same person under variations of physical and mental state and
from person-to-person, thus producing differing levels of variability. The visually ambiguous nature of a leg ulcer is the cause of bias in area measurements and lack of clarity or contrast at the wound boundary adversely affects area measurement precision.

1.4 Proposed Solution

The growth of digital image processing combined with the relative power of the personal computer and the availability of video capture equipment at a reasonable cost has enabled imaging techniques to be extended into many spheres of medical science. It is thus feasible to propose that digital imaging techniques can be used in the process of measurement of wounds. There are many published algorithms that may be applied to the task of segmenting an image and thus attempting to measure the area of a wound. However, it is considered that measurement of wounds by manual delineation is not a trivial task, this being witnessed by the variability of measurements produced by this method. Simple imaging tasks can be achieved by digital imaging methods without the aid of higher-level intervention. However, it is reasonable to consider that an imaging task that is interpreted ambiguously by the human visual system cannot currently be replaced by an autonomous system that is constrained to rely upon the same visual data. Thus an imaging task that is guided by human judgment but which can improve measurement precision is a feasible proposal. To overcome the limitations of the manual delineation process and thus improve upon the precision of wound area measurements, and indirectly, volume measurements, it is proposed that active contour models or 'snakes' (Kass et al., 1987) be used to refine the manual measurement process. Much work involving the development, modification and application of active contour models to imaging specific problems has since ensued. Snakes solve the problem of continuous edge finding in noisy and variable contrast images and, given an initialisation provided by a higher-level process, allow the introduction of a form of globality to the solved edge detection problem. Taking these factors into account, the principal objectives of this project can be stated:
To design or modify suitable active contour model paradigms, tailored to the given application of wound measurement and to determine the appropriate parameter settings with the objective of yielding improved precision of manually delineated area measurements.

To quantify the accuracy of measurements produced by manual wound delineations. Specifically, to estimate levels of delineator bias and variability.

To compare the performance of manual delineation with that of the algorithms proposed. The comparison of performance shall be two-fold:

(a) Quantification of variability: is the area measurement precision improved, or under what circumstances is the precision improved?

(b) Agreement between manual delineator performance and algorithm performance.

This thesis describes the development of four 'flavours' of active contour model algorithms adapting and applying the developments reported in the widespread literature available. The properties of wound images are presented and discussed in Chapter 2 with regard to their bearing upon the application of many image segmentation and edge-detection algorithms reported in the literature, including algorithms applied to other medical imaging problems. It is shown that these algorithms are not easily applied to the wound area measurement problem and that the widely varied properties of different wound images is a great obstacle to the application of many well understood and successfully applied algorithms to this task. Active contour models are chosen as the most applicable solution for several reasons:

- In their native form, active contours are semi-automatic, i.e. they are not necessarily completely automated, requiring an initial estimate of the solution. A manual approximation to the result represents a powerful high-level process input to the problem, although region-growing algorithms could also potentially benefit from such an approach. Ivins and Porril (1994) describe an 'active region model' which
they term a ‘statistical snake’ – a region-growing algorithmic adaptation of the original edge-locating active contour model introduced by Kass et al. (1987).

- Since an approximate initial solution must be given, a final solution almost always exists; i.e. the scope for failure of the algorithm to produce a reasonable result provided it is stable with respect to its initialisation is very limited. There are many images where standard segmentation algorithms (e.g. thresholding, edge following, split-and-merge algorithms) fail to produce any kind of realistic region defining the wound.

- The algorithm does not rely upon homogenous region statistics or uninterrupted boundary contrast, being able to interpolate across poorly defined boundary segments. It is thus more able to cope with the variabilities found both within and between wound images.

1.5 Contribution of this Work

This thesis claims the following as its original contribution to knowledge:

- The tailored design and application of active contour models to the task of wound segmentation for the purposes of area measurement.

- The modelling and analysis of the effect of wound size upon the bias and precision arising from delineated area measurements.

- The analysis of the effect of varying the parameters of the algorithms upon the bias and precision of wound area measurements under manually defined initial conditions.
The selection of a set of working parameters for each algorithm so that each algorithm may be used in clinical practice and produce near-optimised results without the need for human intervention beyond the need to supply an approximate initial solution.

1.6 Definition of Measurement Terms

This section is included to avoid any confusion or ambiguity involving the interpretation of measurement results reported in this thesis. Errors of measurement may be classified as either random or systematic. All experimental error in this work is expressed in terms of precision (as a measure of random error) and bias (as a measure of systematic error). Furthermore, these quantities will usually be expressed with respect to some reference value, e.g. the 'true' area of the wound, giving measurement results of precision and bias in dimensionless per-unit or percentage terms.

**Precision**

Precision is the basic level of variation in a series of repeated measurements. In absolute terms, it shall be defined for a single set of measurements as the standard deviation of the set of measurements. In this thesis calculation of the precision of either manual or algorithmic delineation measurements of wound area is made across a range of wounds. In order to compare the precision of measurements from different wounds it is necessary to express precision as a fraction of the 'true' value, the best estimate of which is regarded as the mean value of the set of measurements. This assumes, however, that a mean value is an unbiased estimate of the true area of a wound. A statement of precision does not indicate whether a bias exists or not. Additionally, the true level of precision is independent of the number of measurements made, meaning that the term 'delineator precision' refers to the expected variability in a single measurement of the area of a wound produced by manual delineation. The definition for fractional precision used in
this work, i.e. standard deviation of a random variable divided by the mean value of the variable, may be identified as the coefficient of variation (Topping, 1972).

Bias

The term bias refers to the difference between the expected value of a wound area estimate and the true area of that wound. This implies that the true area of the wound is known or that a high-accuracy estimate is available. The appraisal of manual delineation performance in Chapter 4 will show that the acquisition of such an estimate is confounded by the fact that different delineators produce differing average area measurements, i.e. their measurements are mutually biased. Mutual bias is thus the measure of the difference between the average values of two sets of measurements, and applies only when the two averages are known to be, or are considered to be, significantly different. This indicates a difference of opinion between two delineators, so that certain parts of the image considered wound tissue by one delineator are not considered so by another delineator and vice-versa.

Accuracy

The term ‘accuracy’ is often used in the context of estimation. As stated above, measurement precision is independent of the number of repeated measurements made of any particular quantity. When many such measurements are made, the resulting estimate is usually expressed as the arithmetic mean measurement and the standard error of this mean result is quoted as the ‘accuracy’. Given the basic level of precision of any piece of measurement equipment, the accuracy or expected error of an estimate made by such equipment may be improved by increasing the number of readings taken. However, standard error does not account for any bias involved in a measurement process, so it cannot be considered a measure of the total expected error unless it is known that no measurement bias exists.
The term *accuracy* will thus be used as a general expression of error when it is appropriate to refer to both types of error in a single phrase. An example of this occurs in Chapter 5 where varying a parameter affects both bias and precision. A graph of bias v. precision is plotted and the path or locus formed as the parameter is varied is referred to as the 'accuracy locus'.
LITERATURE SURVEY

The problem of measuring the size of a leg ulcer given an image may be cast as the equivalent problem of segmenting the wound object from the image background. The immediate background in this case will be the surrounding skin. It is also common for the edge of the limb or foot to be visible so that the limb itself is bordered on either one or both sides by the background to the entire scene. This chapter begins by assessing the image properties required by various segmentation algorithms for good performance and contrasts these with the properties that are common to wound images. The second part of the chapter considers active contour models in detail, considering the problems associated with the original model due to Kass et al. (1987) and the modifications and alternatives reported by later works. Lastly, the chapter discusses the problem of initialisation of active contour models, which if automated could generalise the algorithms proposed in this thesis to be automatic rather than semi-automatic.

2.1 Image Segmentation Methods

The aim of image segmentation is to partition an image into a number of non-overlapping regions that form a complete tessellation of the image plane. A wide range of work has been undertaken to achieve this aim and segmentation has found diverse applications ranging from medical to military uses. It is still a subject of on-going investigation and it cannot be conclusively stated that the segmentation problem has been solved. For the goal of measuring the area of a leg ulcer it is not necessary to obtain a complete segmentation of the image, but to obtain a partial segmentation that discriminates between wounded and non-wounded regions. Surveys of segmentation methods concerning the most common algorithms have been undertaken by Fu and Mui (1981) and Haralick and Shapiro (1985) and most good image processing texts generally
provide a broad over-view of the more common algorithms, e.g. Gonzalez and Woods (1992) and Ballard and Brown (1982). The following discussion appraises the suitability of published algorithms under several sub-headings, although there is inevitably some overlap between them.

2.1.1 Histogram Analysis

The most basic segmentation procedure that may be carried out is thresholding of a gray level image — see Sahoo et al. 1988. The success of this approach hinges upon whether suitable thresholds exist and whether they may be inferred from the image histogram. Thresholds may be applied globally across the image (static threshold) or may be applied locally so that the threshold varies dynamically across the image. Under controlled conditions, if the surface reflectance of the objects or regions to be segmented are uniform and distinct from the background and the scene is evenly illuminated then the resulting image will contain homogenous regions with high-definition boundaries that generally lead to a bimodal or multi-modal histogram. Various methods exist for determining an appropriate threshold, e.g. Weszka (1978), Kohler (1981) and Kittler and Illingworth (1985). For the more general case, where surface reflectances vary and the illumination is not uniform, the threshold needs to vary across the image. An example of a spatially varying threshold based on local histogram analysis is Chow and Kaneko (1972). This algorithm computes a variable threshold function over the whole image by fitting minimum-error thresholds to gray level histograms of local candidate windows.

A novel two-stage threshold process is used by Wu et al. (1995) for segmenting cell images. The authors state that the images used suffer from low contrast and variable intensity and are unevenly illuminated. The region of interest containing the cell is first roughly located by applying a minimum error threshold (Kittler and Illingworth, 1986) to an image of the local gray level variance. A second threshold due to Otsu (1979) is then applied to the local region only, which separates the cell from its immediate
background. Morphological closing and hole-filling operations are subsequently required to remove segmentation artefacts present within the cell region and at its boundary. This appears to be a robust approach to image segmentation that shares some problems common to segmenting leg ulcers. However, the local gray level variance of a leg ulcer image is dependent upon the local curvature of the limb, specular reflections and slough. Figure 2.1 shows example gray level histograms produced from two leg ulcer images.

![Example gray level histograms](image)

Figure 2.1 Example gray level histograms for (a) image 1.1 in PLATE 1. The surrounding skin tissue in this image appears relatively unaffected and is thus quite homogenous. An area of interest was selected that eliminated the background. (b) An example of a wound image (image 2.1 in PLATE 2) which lacks homogeneity of gray scale.

Figure 2.1(a) shows that a shallow valley point exists in the histogram of the image gray level, implying that a threshold could be quite successful in discriminating the wound from the surrounding skin. The images shown in PLATES 1 and 2 are twenty examples extracted from a wound library that has been compiled from hospital visits, a proportion of which have been captured using the MAVIS instrument (Plassmann, 1992). Many wound images show substantial changes in gray level as a result of practical difficulties that arise in obtaining well-illuminated images. Difficulty in evenly illuminating highly curved parts of the leg is a common cause of this problem. Clearly, this problem is exacerbated as the physical size of the wound increases. Further difficulties exist in accurately locating the wound boundary when epithelialisation tissue is present at the
wound edge, examples of which are images 1.3, 1.4 and 1.5 in PLATE 1 and image 2.7 in PLATE 2.

2.1.2 Region Growing and Split-and-Merge Algorithms

Region growing algorithms start with a number of seed pixels or seed regions and grow these regions by adding to a region previously unassigned neighbouring pixels that correspond to some similarity criterion for that region (Gonzalez and Woods, 1992). The rule employed for accepting/rejecting a candidate pixel for merging with the neighbouring region may be defined as either on the basis of a pixel and its immediate neighbours or on its relationship to the mean of a region (using either Euclidean or Mahalanobis distance as the discriminant). Growing algorithms that link neighbouring pixels by analysing contrast use edge-strength to measure the compatibility between neighbouring pixels (edge strength) and thus require a completely closed boundary of high-edge strength pixels, otherwise neighbouring regions will be merged. Growing algorithms that measure similarity by comparing a pixel with some measure of the whole region (e.g. mean value) are more robust but impose a stronger homogeneity criterion upon the regions being grown. Recently, Adams and Bischof (1994) presented a region-growing algorithm that is robust and free of tuning parameters. However, they state that the algorithm requires suitable pre-processing to remove background variation due to uneven illumination. Another recent development of the region-growing paradigm is the 'statistical snake' due to Ivins and Porril (1994). This algorithm imposes a smoothness constraint on the boundary of a region. Snakes (active contour models) are discussed in §2.2.

Split-and-merge algorithms have been reported, among others, by Klinger (1973) and Horowitz and Pavlidis (1976). This method proceeds to successively divide an image into smaller non-overlapping regions if some similarity criterion is not met, otherwise no split of that region is made. The end result of the splitting is an over-segmented image. A merging procedure is then applied to merge neighbouring regions under the
same homogeneity predicate that was used for splitting. Homogeneity tests may be based on simple image statistics, e.g. the range of intensity values in a region, or upon more robust statistical similarity measures such as the Kolmogorov-Smirnov test (Muerle and Allen, 1968). Chou et al. (1992) use an adaptive split-and-merge method to segment magnetic resonance (MR) and x-ray computer-tomography (CT) images for calculation of volume and reconstruction of internal organs. They use a likelihood ratio test based on a Gaussian distribution assumption to check for region homogeneity. The lack of homogeneity across the wound in many images of leg ulcers and the variable contrast make split-and-merge schemes unsuitable for measuring the area of many wounds (see PLATES 1 and 2 for examples of leg ulcer images).

2.1.3 Colour-Space Analysis and Segmentation

Recursive region-splitting algorithms using colour-based discriminants have been described by Ohlander et al. (1978) and Ohta et al. (1980). Starting with the whole image as one region, these algorithms apply a set of thresholds derived from analysing histograms of individual colour features. Once a region has been segmented, the algorithm is recursively applied to each new region. Ohta et al. examined a range of images and determined a set of three colour features which had near-optimal discriminant power. The primary discriminant in the cases they analysed was found to be equal to the gray level or intensity component. Arnqvist et al. (1988) perform colour analysis of wounds using Red/Yellow/Black classification, corresponding to granulation tissue, slough and necrotic material respectively. Colour images of leg ulcers are automatically analysed by Beriss and Sangwine (1997) using a technique based on 3-D histogram clustering of the colour values. Clustering of the histogram is achieved by erosion and dilation operations that break the connections between the clusters corresponding to different tissue classes. The result of segmenting one wound image into regions of granulation tissue, slough and surrounding skin is presented.
Umbaugh *et al.* (1993) present two colour segmentation algorithms, which they use to segment images of skin tumours. The first algorithm represents the chromatic component of an image by two angles representing the orientation and elevation of a colour vector in the RGB colour space. This is equivalent to normalising the colour vector using the $L^2$-norm (Healey, 1992). The second algorithm recursively applies a principal component transform (PCT, also known as an eigenvector transform or the discrete Karhunen-Loeve transform (KLT)) to the colour data followed by a split of the colour space at the median point along the principal axis. This process is repeated on each subspace so formed until the colour space is divided into the desired number of colours. A segmentation based on global colour properties as produced by an RGB colour histogram assumed that the colour properties of the wound are homogenous under the combination of illuminants used and level shifts and gamma-transformations applied by the camera and image-capture equipment.

Schmid and Fischer (1997) segment pigmented skin lesions using a principal component transform of RGB colour images of digitised epiluminescence microscopy slides. The 2-D histogram of the first and second principal axes of the data is smoothed by a morphological process before being clustered into three classes of tissue using the fuzzy $k$-means algorithm. One example of the smoothed 2-D histogram is presented showing good cluster separation. The segmentation result is affected by hair and the inherent texture of the lesion and so a median filter is applied before image segmentation and the subsequent boundaries detected are smoothed and small regions removed by morphological opening and closing.

### 2.1.4 Pixel Classification for Segmentation

Image segmentation can be achieved by grouping together neighbouring pixels that are similarly classified. The classification may be based on trained parameters (supervised classification) or based on clustering techniques, e.g. the $k$-means and ISODATA algorithms (Duda and Hart, 1973). The automated colour-image analysis mentioned
above is an example of unsupervised classification or clustering. Classification schemes can be potentially improved by using contextual information, basic examples of which are relaxation labelling (James, 1987) and Markov random field methods (Therrien, 1989). Däschlein et al. (1994) used an artificial neural network (ANN) to perform the classification task for segmenting CT images. The features assigned to each pixel were based on five basic measures of a local co-occurrence matrix. A contextual classification of medical images is reported by Ossen et al. (1994). They use an ANN to estimate the a-posteriori Bayesian probabilities for a minimum-error classification to aid in the diagnosis of Graves' ophthalmopathy. Jackson et al. (1993) compared the performance of a relaxation-labelling algorithm with a minimum Mahalanobis distance classifier to segment MR and CT images. They report that the relaxation scheme is more consistent in labelling regions of gray and white matter and cerebro-spinal fluid (CSF). Whilst the visual appearance of the results is impressive it is necessary to consider that the regions of the 'gray level' image corresponding to particle densities are somewhat distinct and homogenous. This is in contrast to the conditions found in many images of leg-ulcers.

2.1.5 Gradient Operators, Edge and Zero-Crossing Detectors

Canny (1986) proposed the derivative-of-Gaussian filter as a near-optimal filter with respect to three edge-finding criteria: (a) good localisation of the edge, (b) one response to one edge and (c) high probability of detecting true edge points and low probability of falsely detecting non-edge points. Deriche (1987) implements a filter with an impulse response similar to that of the derivative of Gaussian, but which lends itself to direct implementation as a recursive filter. This requires fewer coefficients than a transversal or non-recursive filter that has the same impulse response, yielding faster convolution, especially when the width of the filter becomes large. The benefit of these filters is that they impose a smoothing upon the image and as such are band-pass filters rather than simple differentiators and thus they are less sensitive to noise. Smoothing is controlled by a scale parameter which may be tuned for each application. Edges are extracted from
the gradient image by choosing pixels with a gradient magnitude that is a maximum in the local direction of the gradient.

Marr and Hildreth (1980) have proposed finding zero crossings produced by the Laplacian-of-Gaussian (LoG) operator for the detection of edges. This operator is orientation independent and accurately locates zero crossings, provided that the smoothed gray level variation along, but not necessarily near to, a line of zero-crossings is linear. Haralick (1982) proposed fitting bi-quadratic facet models to an image and applied a zero crossing detector to the second derivative of the function taken in the direction of maximum gradient. This is a more computationally expensive method than the LoG operator proposed by Marr and Hildreth.

In relation to finding the edges of wounds, the application of such derivative operators generally fails to produce closed boundaries – this applies both to maximum gradient detectors and zero crossing detectors. Although it is possible to close small gaps by extending the ends of detected edge chains (Niblack, 1987) it is quite common for edge segments to curve into or away from the wound boundary, even at high levels of smoothing, so that one end of the detected edge segment is positioned some distance away from the wound boundary. In Figure 2.2 the detected edge contour deviates outside of the wound at the lower left side of the wound. This image is shown in colour in PLATE 2 (image 2.7). However, for a broad range of leg ulcers, the information provided by a gradient operator provides a strong cue for detecting the edge of a wound. Clearly, such operators will detect all edges and not just those belonging to a wound.
Edge focusing algorithms (Bergholm, 1987) use smoothed gradient operators and detect and track edges throughout a scale hierarchy. There is a much higher likelihood of locating the boundary once the gradient image has been smoothed to a relatively high degree. Typically, the smoothing parameter is reduced by a small amount so that the boundary moves by only a small amount (e.g. 1 pixel), allowing the algorithm to accurately track the movement of the contour detected at the previous scale. Focusing (de-blurring) allows the contour to converge towards its true location in the original unblurred image, since the smoothing displaces and smooths apparent contours in the image. Denton et al. (1995) use an edge focusing algorithm to detect skin lesion boundaries. Accurate location of the boundary is important so that automated diagnosis can subsequently be made using colour, texture and shape measures to classify the lesion, which may be cancerous. The approximate size of the lesion is assessed by an automated bi-level thresholding operation (Kittler et al., 1984) after pre-processing by a median filter. The estimated lesion size is used to tune the initial scale-constant of the LoG filter used for detecting lesion edges. The success of edge focusing relies upon detecting most of the required boundary at high scales. If any gaps in the boundary were to be left unlinked then neighbouring ends of the edge chains could migrate as the image was de-blurred. Such edge chains occur infrequently in wound images thus limiting their applicability within the scope of leg ulcer measurement.

Perona and Malik (1990) propose a multi-scale or scale space representation of an image based on anisotropic diffusion. This models the image smoothing as the diffusion of heat energy in a physical system. The heat conduction coefficient is allowed to vary across the image plane and is made dependent upon the image gradient. This effectively leads to a spatially adaptive smoothing which tends to preserve the location of edges throughout the scale hierarchy. The idea behind this method is that it avoids unconstrained smoothing of the image across edges, which causes the edges to move and thus prevents them from being accurately located at high scales. The problem formulation leads to an iterative minimisation algorithm and thus the algorithm imposes an increased computational cost.
2.1.6 Watershed Segmentation

The watershed algorithm may be applied to a gradient magnitude image for the purposes of image segmentation (Haris et al., 1998 and Gauch 1999). Considering the gradient magnitude at each pixel as the height of a surface in 3-D, regions are formed by simulating a flooding of the image that begins at local minima of the gradient image function. The boundaries of each region stop advancing when neighbouring flooding regions meet and the boundaries thus correspond to ridge-paths of the gradient magnitude. The principal advantage of the watershed segmentation scheme over other edge-finding schemes is that it generates closed boundaries. The regions defined by the closed boundaries represent an oversegmentation of the image, since there will usually be many such regions corresponding to each object. If the gradients are computed at successively higher scales, the number of local minima (flood basins) in the gradient magnitude image will decrease. Watershed segmentation can be applied to wound images and will form a sensible oversegmentation of many of the moderate and well-defined images. Figure 2.3 shows an example of applying the watershed segmentation algorithm to a derivative of Gaussian gradient magnitude image. Although the watershed algorithm can produce boundary segments that correspond to each part of the true wound boundary (approximately), the problem of merging wound regions into one region is confounded by the fact that very often two regions, one on either side of the boundary, are similar in their gray level properties and are very likely to be merged. A possible improvement in comparing regions for merging is to make use of second order image statistics in the form of a co-occurrence matrix, as proposed by Chen and Pavlidis (1979).
2.1.7 Radial Search Algorithms

Radial search algorithms simplify the task of edge detection by constraining the search for an edge point to lie on a line emanating radially from a central point within the object. The method also simplifies the detection of erroneously marked points by allowing a straightforward comparison between edge points marked on neighbouring radial lines. Golston et al. (1990) used a radial search algorithm to detect skin tumour boundaries having considered that standard methods were only suitable for a small subset of tumour images. A heuristic measure for edge-worthiness along each radial search line requires the luminance to increase by a specified minimum amount for a specified minimum sustained number of pixels. The authors reported good results for 17 out of 20 tumour images tested. Radial search algorithms are only applicable to delineating objects that are mainly convex. Ruiz and Fairhurst (1995) detected cardiac boundaries in 2-D echocardiographic images by applying a LoG filter and zero-crossing detector to the image prior to radial sampling. O'Gorman et al. (1983) use a radial search to segment cell nuclei. On each radial line a sigmoidal-shaped monotonically increasing edge model is used to select an appropriate gray level threshold. Although such algorithms have been applied to difficult segmentation tasks with low and variable boundary contrast, leg ulcers give rise to many potential edge-point candidates within the wound itself.

2.1.8 Boundary Following Algorithms

Boundary following algorithms have been used to find the edges of cells in microscopy images. An algorithm due to Jarkrans et al. (1980) segments cell nuclei and cytoplasm by searching for gradient maxima along 8 radial lines emanating from a manually selected seed point. The points detected by the radial search are joined by a tracking procedure that attempts to follow a path of maximal gradient given by a chain-code of the gradient direction. Not all attempts to track cell-boundaries were successful in producing fully closed boundaries. Graph-searching techniques for boundary detection
based on a heuristic graph-searching algorithm by Nilsson (1971) have been reported by Martelli (1972) and Lester et al. (1978).

The principal disadvantage of boundary-following algorithms is that enough boundary information needs to be available in the image to form an acceptable path between start and end search nodes. Expanding all possible search nodes between two points does not guarantee that a path will be found linking the two. As already stated, the gradients determined for leg ulcer boundaries form incomplete paths, are ambiguous and can trace their way into the wound rather than following the actual boundary.

2.1.9 Summary

This section has discussed various methods of image segmentation and some of their applications. Many of the methods could be used to segment some of the images that are more evenly illuminated, free from specular reflection and more homogenous with regard to their gray level properties. However, such wounds are fairly uncommon among the wound images observed and in general any segmentation algorithm will produce progressively less satisfactory results as the ideal conditions for an image degrade. The vast variability in the quality of wound images with respect to these idealised characteristics generally prevents any one algorithm from being universally applied. If an algorithm is to be successfully applied to measuring the area of leg ulcers then it has to be capable of producing sensible results on the majority of wound images. Clearly, the accuracy of such an algorithm will be affected by the qualities of the wound image. However, if the algorithm can still produce a reasonable measurement result when the image quality is poor then this represents a best solution under the circumstances. It is this variability which precludes the general use of any algorithm which relies upon some quasi-homogeneous property to distinguish the wound from the surrounding tissue. Most of the images of leg ulcers in the library show variation in boundary contrast.
2.2 Active Contour Models

Active contour models (also known as ‘Snakes’) were introduced by Kass et al. (1987) as a novel solution to the low-level imaging task of finding salient contours (e.g. edges and lines) in digitised images. An interesting and powerful property of an active contour model is its ability to find subjective contours and interpolate across gaps in edge chains. An active contour model represents an object boundary or some other salient one dimensional image feature as a parametric curve that is allowed to deform from some arbitrary initial position towards the desired final contour. The problem of finding this final contour is cast as an energy minimisation problem with the intention that it yields a local minimum of the associated energy functional. The energy functional is thus based upon the salient features of the image under inspection so that when the contour is aligned with such a feature, the energy of that contour corresponds to a local energy minimum. The minimisation of the energy functional gives rise to spatial forces that act upon the contour to deform it by translating, stretching and bending it as necessary. It is necessary to include in the energy functional terms that impose smoothing constraints upon the contour since the image itself is corrupted by noise from various sources. The original model incorporates two internal energy terms acting to constrain the solution by imposing conditions related to contour smoothness and regularity. These properties are discussed in §2.2.3. Implementation of an active contour model involves several basic considerations that are explored. All aspects of the original model have been subject to modifications and extensions by later researchers that have served to address perceived weaknesses of the complete implementation.

2.2.1 Modelling of the Contour

An active contour model algorithm comprises the fitting of a deformable model of either a closed-boundary or open-boundary-segment to available image data, e.g. the outline of an object and in the case of the present work, the edge of a wound. The contour is defined in the \((x, y)\) plane of an image as a parametric curve, \(v(s) = (x(s), y(s))\) where \(s\)
is a parameter advancing around/along the contour and is related to arc length. This is a continuous-domain representation of the contour. Subsequently it will be necessary to represent the curve in a discretised form where \( v(s) \) is either (a) uniformly point-sampled in the domain of \( s \) (finite difference discretisation) or (b) represented by a sequence of piecewise polynomial segments (finite element discretisation). It is desirable in this instance that the curve possesses continuous first or second order derivatives across the boundaries of neighbouring elements – referred to as \( C^1 \) and \( C^2 \) continuity respectively. The discretisation step is considered in §2.2.6.

### 2.2.2 Formulation of the Energy Functional

Having specified the contour as \( v(s) \), the deformable model is defined as a sum of energy terms in the continuous spatial domain, where the energy terms may be categorised as (a) internal, i.e. functions of the contour \( v(s) \) itself, or (b) external, generally functions derived from the image under inspection. The original model defines a third category, constraint energy, consisting of higher-level factors added to provide more control over the contour than the internal energy terms, although such factors have not been used in this work. The continuous-domain energy functional for the Kass snake with one external energy term and without constraint terms is given by:

\[
E_{\text{snake}}^* = \int_0^s \left( \frac{\alpha}{2} |v_s|^2 + \frac{\beta}{2} |v_{ss}|^2 + P(v) \right) ds
\]

where \( \alpha \) and \( \beta \) are coefficients governing the elasticity and stiffness of the contour \( v(s) \) respectively and \( P(v) \) is an external energy (potential) function sampled along the contour. The convention is adopted of denoting a derivative or partial derivative of a function by appending the appropriate subscript, e.g. \( v_{ss} = \frac{d^2 v(s)}{ds^2} \).
2.2.3 Minimisation of the Energy Functional

Generally the potential function $P(x,y)$ will give rise to many local minima of (2.1). Any contour $v^*(s)$ that corresponds to a local minimum of (2.1) is a solution to the contour-finding problem. Associated with such a functional is a set of governing differential equations, termed the *Euler-Lagrange* equations, the solution of which yields the energy-minimising contour. Farlow (1982) gives a good introduction to the task of deriving such equations. Minimisation of the energy model of (2.1) gives rise to the following pair of differential equations, the terms of which equate to the various internal and external forces that are balanced when (2.1) is locally minimised:

\[
-\alpha v^*_{ss} + \beta v^*_{ssss} + \nabla P(v^*) = 0
\]

where $v^*(s)$ is a local minimiser of (2.1)

Substituting an arbitrary contour $v(s)$ for the minimising contour, $v^*(s)$ into (2.2) generally yields a non-zero residual term. It is this residual function which is responsible for deforming a contour $v(s)$ towards the minimising contour $v^*(s)$.

It will help to elucidate the nature of the component energy terms and respective forces by way of some elementary analyses.

2.2.3.1 Elastic Energy and Forces

The elastic energy component of the active contour energy functional (2.1) and its resultant forces taken from the Euler-Lagrange equations (2.2) are given respectively by:

\[
E^*_{elastic} = \frac{\alpha}{2} \int |v_s|^2 ds
\]

\[
F^*_{elastic} = -\alpha v_{ss}
\]
The energy term (2.3) is minimised (to zero) when the first-derivative of the contour is identically zero everywhere, signifying that the contour of minimum elastic energy is one that has collapsed to a point. One may wish to specify end conditions for a contour which can be of open or closed (cyclic) form. Consider the case that arises when the contour is an open curve with fixed end points (hard constraints upon (2.1)): clearly the constraints prevent the contour from collapsing to a point and the action of the elastic forces is to straighten the curve. Such action requires forces that are generally directed into the centre of curvature. When the contour is closed the action of the elastic forces is unchanged, although ultimately the contour is deformed to a completely convex object and the forces become directed towards the inside of the object. Figure 2.4 shows a parametrically-defined closed contour upon which are superimposed the force vectors resulting from the second derivative (elastic) term of (2.3). Since the force vectors are always oriented in the direction of the centre of curvature at each point on the contour, at a concavity the forces are oriented outwards from the inside of the object. This causes the contour at these points to expand until the concavity is eliminated. The force vectors in Figure 2.4 are not normal to the boundary however, indicating that the movement of the boundary into the direction of curvature is only one mode of movement. The secondary action of elastic forces is shrink the curve – this action is tangential to the contour and in the case of the open curve is the force component responsible for collapsing it to a point in the absence of competing or constraining forces. The decomposition of elastic force into its normal and tangential components is considered in (§3.2).
2.2.3.2 Bending Energy and Forces

The bending energy term and corresponding force term of a contour \( v(s) \) are given respectively by the following:

\[
E_{\text{bending}}^i = \frac{\beta}{2} \int |v_{ss}|^2 ds
\]

\[
F_{\text{bending}} = \beta v_{ssss}
\]

Sharp corners or points of high curvature and fine contour detail are characterised by high frequencies. In comparison with the elastic energy term, the bending energy term is more sensitive to these higher-frequency components because it is based on a higher-order derivative. Figure 2.5 shows a comparison between two open contours that are constrained to pass through 3 points. Under this constraint, the contours are in equilibrium (energy minimised) for (a) an elastic energy term only and (b) bending energy term only. At the central point in Figure 2.5 the contour's first derivative is discontinuous and its second derivative is thus infinite. The distribution of the bending energy of such a contour would be infinite and also completely concentrated at this one point. The lack of high curvature and sharp corners implies low bending energy and the curve is smooth and \( C^1 \) continuous everywhere. However, as with the elastic energy term, (2.5) is scale-dependent and the bending energy of a contour can always be reduced, not just by attempting to distribute the curvature evenly along the contour, but also by a contraction of the whole contour. Without opposing forces, the bending energy term will cause a contour to naturally degenerate to a

![Stiff Contour in Equilibrium](image)

![Elastic Contour in Equilibrium](image)

**Figure 2.5** Two minimum energy contours constrained to pass through a set of three arbitrary points (marked by circles).
straight line. This form of (2.4) in fact dictates that both $x(s)$ and $y(s)$ components of $v(s)$ both be linear, rather than simply being a linearly related arbitrary form. When the contour is closed the natural degeneracy of the contour under the minimisation of (2.4) is a point since the contour is constrained from becoming linear. This form of degeneracy is a manifestation of the scale-dependent nature of the energy term.

2.2.4 Weaknesses of the Kass Model

The original active contour model suffers from the following weaknesses:

- The internal forces have the side effect of tending to shrink open contours (unless the end-points are firmly anchored) and can cause closed contours to collapse to a point. This makes setting of the weights problematic.
- The particular form of the external energy tends to produce bunched-up points — see §2.2.10.
- The solution is initialisation-dependent. Image noise may give rise to spurious local energy minima that trap the contour before it reaches the desired solution. Also, an external energy function based on edge-strength alone will cause the contour to be attracted to any such nearby edge, regardless of whether such an edge belongs to the desired solution.
- The time-stepping evolution equations that are used to iterate the initial approximation towards the final solution are only conditionally stable. This imposes constraints upon the coarseness of the time-step and thus slows convergence.
- Spatial discretisation of the model using finite differences requires many points to adequately represent the contour. This leads to large matrices requiring much computational effort to invert if one uses the semi-implicit iteration method to promote stability of the solution.
Several authors have proposed various methods of overcoming these difficulties and such modifications and improvements are considered in the following section. However, it must be stated that none of the modifications succeed in overcoming the lack of globality of the solution, requiring good initialisation in order to produce a satisfactory solution.

### 2.2.5 Solution of Euler-Lagrange Equations

When the active contour model is not in a stable equilibrium state (corresponding to an energy minimum) the Euler-Lagrange equations of (2.2) result in a set of unbalanced forces so that the right-hand side (RHS) will not be zero. The solution may be sought by setting the RHS to a negative time-varying residual force term. This is similar in operation to dissipating some of the energy of the system via a frictional or resistive component and is the method that was used by Kass et al. (1987). A physical interpretation of the system has been provided by Terzopoulos (1987) that includes an inertial (mass) term allowing a full physical interpretation of the model including the dynamics of the solution. The inertial component allows the model to overcome the attractive forces acting in the shallow energy wells of undesirable and noisy ‘solutions’. In the continuous time and space domains the full inertial model for system motion may be defined by:

\[
(2.7) \quad -\alpha v_{ss} + \beta v_{ssss} + F(v) = \mu v_{tt} + \gamma v_t
\]

where \(\mu\) is the mass density of the contour and \(\gamma\) is a damping constant.

This system is in a stable equilibrium state when the time-derivative dependent terms of the RHS decay to zero so that (2.7) corresponds to (2.2). Equation (2.7) may be suitably discretised in both temporal and spatial domains and subsequently solved by iteration through discrete time. Setting \(\mu\) to zero in (2.7) yields the original solution form that corresponds to a mass-less contour which thus lacks the ability to overcome weak local
minima. Now, any physical system which stores energy in more than one medium and which is not at rest has the potential to undergo oscillations, where energy is cyclically transferred from one storage medium to another. A simple example is a swinging pendulum where the pendulum’s energy is cyclic exchange of potential and kinetic energy. Setting $\gamma$ to zero in (2.7) corresponds to a non-dissipative or Hamiltonian system and in this case any energy ‘lost’ by virtue of the action of the forces on the left-hand side (LHS) is transferred to the inertial term. None of this energy is dissipated or irrecoverably lost from the system, hence the total energy of the closed system (no energy transfer either in or out of the system) is a constant. Thus, even if the contour $v$ corresponded exactly to the desired energy minimum – both sides of (2.7) being zero – the system would not be at rest. For the system to stabilise and converge to a unique energy minimum it must lose energy by dissipating it via the frictional term. A physical analogue of the original evolutionary equation may be made by considering the mass-density parameter $\mu$ to be very small. Note that letting $\mu=0$ implies that $\gamma=0$, otherwise the physical analogue breaks down. The Rayleigh dissipation constant $\gamma$ plays an important role in the stability of the system. If $\gamma>>\mu$ the system is clearly over-damped and in theory should converge to an energy minimum with no overshoot or oscillation. However, the complex topology of the external potential function $P(x,y)$ dictates that the equations must be solved numerically by discrete time stepping. Transforming the continuous-time system of (2.7) into a discrete time system introduces the possibility of instability.

2.2.6 Spatial Discretisation

The differential equation of (2.7) may be spatially discretised by representing the contour $v(s)$ and its derivatives using either finite differences or finite elements. The matrix coefficients arising from finite differencing of (2.7) are given by Kass et al. (1987). In this application it is proposed that finite element discretisation be used. The application of the finite element (FE) method for solving differential equations to the problem of discretising and solving the continuous-domain energy formulation of active
contour models has been reported by Karaolani et al. (1990) and Cohen and Cohen (1993). Karaolani et al. compare the use of parabolic and hermite cubic polynomials for spatially discretising an active contour model. The derivatives of a parabolic curve vanish after the third derivative making them unsuitable for active contour models requiring a bending energy component. It is possible, by use of the Rayleigh-Ritz method or by judicious use of integration by parts, to reduce to two the number of derivatives required by a finite element implementation that contains a bending energy term. However, cubic elements provide a method for applying C\(^1\) continuity to a curve, which is a desirable aspect when the curve is required to be smooth. They also allow more flexibility in their application.

The finite element approach has the following advantages over finite difference schemes:

- FE requires fewer elements than FD requires points, thus the matrices are much smaller and hence allow for faster iteration. Also, the elements are then generally much longer than the distance between sample points in a FD scheme and thus there is less likelihood of the contour intersecting itself.
- The sampling of the image (potential function) is much finer when using a suitable number of quadrature points for evaluating the required integrals. This allows the model to become more sensitive to fine detail in the potential function and less sensitive to noise.

2.2.6.1 Finite Element Shape Functions

The most convenient form of interpolating curve for this application is that of piecewise-polynomial curves, provided they possess at least C\(^1\)-continuity to produce smooth boundaries, are twice differentiable and that changes to any one of the control vertices only affects the piece-wise curve locally. Both cubic Hermite and cubic B-Spline curves possess these qualities (Bartels et al., 1987) and are readily used in the
formulation of a finite element active contour model, preferably with some adjustment in the case of Hermite curves. Discretising the parametric curve \( v(s) \) into finite elements leads to an expression for each element given by:

\[
(2.8) \quad v^e(s) = N(s)V^e
\]

where \( N(s) \) is the set of shape or basis functions defining the form of the curve and \( V^e \) is a two-column matrix of control vertices.

The use of both Hermite cubic and cubic B-Spline elements is considered next. Definitions of the respective shape functions are given in Appendix A.

**Cubic Hermite Basis**

Interpolation with a cubic Hermite basis requires specification of both end-points of the curve and both end-derivatives of the curve. This means that the matrix equation for the curve is not directly expressed in the form required for application with the finite-element method. Such curves have great application in the field of computer graphics and are particularly useful for producing inter-frame interpolation for animated sequences (Bartels et al., 1987). In such applications the need to specify the end-point derivatives of the curve segments is a powerful tool for applying tension to the interpolating curve and also allows a bias to be added so that derivatives may be skewed. Furthermore, one may desire to break the implicit \( C^1 \)-continuity at the knot points between piece-wise curve segments, so producing a corner, by allowing the two derivative values at this point to differ. Thus, in general, there is good reason to retain control over the end-point derivatives. Hermite curves have the property of interpolating their control vertices.
Cubic B-Spline Basis

The set of cubic B-Splines has the property of producing $C^2$-continuity between curve segments (elements). It should be noted that cubic B-Splines do not interpolate the control vertices. To fit a unit-spaced B-Spline that interpolates a set of element end-points $P^e$ one must choose control vertices $V^e$ such that:

\[
(2.9) \quad P^e = B(s = 0)V^e = \frac{1}{6} V^{e-1} + \frac{2}{3} V^e + \frac{1}{6} V^{e+1}
\]

Thus, given an approximate wound boundary suitably represented by a vector of element end-points $P^e$ one must obtain the control vertices $V^e$ that interpolate the end points and perform iterations of the active contour model until convergence is attained. Subsequently, the final element end-points may be calculated from the $V^e$ that represent the converged contour. The matrix of equations defined by (2.9) is tri-diagonal and diagonally dominant, allowing rapid solution.

Comparison of Hermites and B-Splines

Figure 2.6 shows a comparative plot of the cubic Hermite and cubic B-Spline shape function sets, produced by plotting the respective shape equations listed in Appendix A. The use of cubic B-Splines gives rise to smoother curves (by virtue of higher-order continuity between elements), but such a curve does not interpolate its control vertices, as stated above. The higher order continuity between the elements of a cubic B-Spline curve tends to produce smoother curves than those defined by cubic Hermite elements.
2.2.6.2 Stiffness Matrix Equation

The discretisation of the static problem given equivalently by (2.1) and (2.2) yields a matrix equation. At a unique and stable equilibrium, the complete discretised system of equations may be written in matrix form as:

\[(2.10) \quad KV^* + F(V^*) = 0\]

where \(V^*\) represents the control vertices defining the curve of specified form which locally minimises (2.1).

The derivation of this form of equation from (2.1) is given in Appendix A for both cubic hermite and cubic B-Spline elements. The size of the \(K\) matrix, its sparseness and the precise values of its coefficients are dependent upon the number of points or elements (and their form) used to represent \(v(s)\) discretely. The cubic hermite curve gives rise to coefficients of significantly larger magnitude than those of the B-Spline, particularly in regard to the stiffness terms. Leymarie and Levine (1993) show that the conditioning of
the matrix equations is partially dependent upon the size of the eigenvectors of $K$ which in turn are affected by the magnitude of the coefficients. However, this is unlikely to be a matter of concern unless high values of $\beta$ are used and the equations are solved with large time steps.

2.2.7 Time Discretisation

A direct single-step analytical solution of (2.10) for $V^*$ is not feasible since the external forces which give rise to the $F(V)$ term are non-linear. Therefore the aforementioned discrete time stepping method of solution is necessary. When the system of equations is not in equilibrium, the RHS of (2.10) is non-zero and thus (2.10) may be generalised to a time-dependent version given by:

\[
KV_{(n+1)T} + F(V_{nT}) = \frac{T}{\gamma} (V_{(n+1)T} - V_{nT})
\]

Rearranging (2.11) for $V_{(n+1)T}$ yields an iterative equation:

\[
V_{(n+1)T} = \left( I + \frac{K T}{\gamma} \right)^{-1} \left( V_{nT} - \frac{T}{\gamma} F(V_{nT}) \right)
\]

Equations (2.11) and (2.12) are essentially the equations used by Kass et al. (1987). The internal forces are linear and specified wholly by the stiffness matrix $K$ allowing implicit or \textit{backward} Euler steps to be taken with respect to these forces. Explicit or \textit{forward} Euler steps are taken with respect to the external energy since its form is generally non-linear and cannot therefore be easily expressed by an analytical expression at time $t=(n+1)T$. The external forces are thus treated as remaining constant during each time-step. This mixture of implicit and explicit stepping is loosely termed 'semi-implicit'.
Chapter 2 Literature Survey

The solution of (2.12) at each iteration may be implemented in several ways. Kass et al. (1987) and Cohen (1991) use LU decompositions, while Leymarie and Levine (1993) suggest using either Cholesky decomposition or a penta-diagonal algorithm due to Benson and Evans (1977). Cholesky decomposition is applicable when the matrix to be 'inverted' is both symmetric and positive definite (all positive eigenvalues). Press et al. (1992) state that "Cholesky decomposition is about a factor of two faster than alternative methods for solving linear equations." Further considerations for speed improvement focus upon exploiting the sparseness of the stiffness matrix – see Benson and Evans (1977) – and upon how often the matrix changes (e.g. by refining $T$) and thus how often it requires re-inverting. If the matrix remains constant its inverse may be calculated from any of the above decomposition methods and stored so that subsequent steps require only matrix addition and multiplication. In the present case however, such considerations need not be a matter of great concern since (a) a wound does not require a large number of elements to define its boundary and (b) execution time of a wound-measurement algorithm is not too critical. The MAVIS instrument produces results within five minutes of the wound image being taken (Plassmann, 1992).

2.2.8 Stability of the Iterative Solution Method

Numerical solution of equations by implicit stepping methods is only guaranteed to be stable when a system is linear. However, the external force term in (2.12) gives rise to possible numerical instability. If the external forces dominate the internal contour smoothing forces in (2.12) and too large a time-step is taken, the contour could easily be moved out of the domain of influence of a local edge-profile and become influenced to converge to another edge that is not part of the desired solution. Cohen (1991) addresses this situation by normalising the external forces to have unit magnitude and adds a scaling factor, $k$. The external forces then become $k\tau F_i/|F_i|$ where $\tau = T/\gamma$. A second problem Cohen considers occurs when a contour attempts to stabilise in alignment with a path of maximum gradient. Since the external force function is defined on a finite grid, the forces are suddenly reversed as the contour moves under the action of those forces.
from one side of the edge to the other. Figure 2.7 shows how this can result in a limit cycle oscillation:

![Figure 2.7](image)

(a) Production of limit cycles from non-interpolated external forces

(b) Convergence of steps due to use of linear interpolation

However, this does not complete the analysis since the smoothing applied by \((I + \tau K)^{-1}\) can itself cause the contour to move by a large amount when the internal forces dominate the external forces – this is independent of the choice of \(\tau\). Equation (2.13) analyses the problem by expressing the displacement of the control vertices under the action of (2.12):

\[
(2.13) \quad \Delta V_t = (V_{t+1} - V_t) = -\tau (I + \tau K)^{-1} \left[ K V_t + \frac{k F_t}{|F_t|} \right]
\]

The term in the square brackets is easily recognisable as the residual force term from setting (2.10) to be non-zero, with the external forces modified as suggested by Cohen. Clearly, both parts of the residual at time \(t\) are smoothed by the inverse-matrix term and scaled by the time step to produce the displacement of the control vertices controlling \(v(s)\). Clearly, when \(\alpha\) and \(\beta\) are large enough the internal forces dominate the equation – but there is no fixed limit over exactly how far the time-step taken will cause vertices and hence the contour to move. The displacement of the control vertices given by (2.13) when the internal forces are dominant is a function of \(V_t\) itself, \(\alpha\) and \(\beta\) as well as \(\tau\).
Therefore, one cannot ensure that a contour will definitely behave in a steepest descent manner by adjusting the time step alone.

2.2.9 Direct Energy Minimisation Approaches

Several authors have proposed iterative methods of directly minimising energy functionals of similar form to (2.1). Amini et al. (1991) employed a Dynamic Programming technique (Ballard and Brown, 1982). This involves evaluating the contour energies produced by local changes to the nodes of the contour and is numerically stable and guaranteed to converge to a unique solution within a local minimum of the energy functional. This approach allows the setting of hard constraints and Amini cites the example of constraining the spacing between contour nodes to be greater than a certain specified distance. However, the dynamic programming method of solution has two main disadvantages: (1) It is relatively slow, executing in $O(nm^2)$ time where $n$ is the number of contour sample points (‘snaxels’) and $m$ is the number of trial offset patterns associated with each snaxel; (2) the energy functional must be capable of being decomposed into locally dependent terms so that the energy associated with each snaxel is affected by only a few neighbouring snaxels.

To overcome these disadvantages, Williams and Shah (1992) have proposed a ‘greedy’ algorithm, which considers the energy of a single snaxel at each step and moves it to the neighbouring position that minimises the associated energy term. The greedy algorithm executes in $O(nm)$ time, sacrificing the ability to converge completely to a stable minimum. Lai and Chin (1993) propose a modified greedy algorithm that is guaranteed to converge. This is achieved by minimising of the sum of all the local energy terms that are affected by movement of any single snaxel. The original active contour model uses a semi-implicit scheme to promote stability under all settings of the step size parameter, given some upper limit on the size of the external forces. The trade-off for efficient convergence requires a balance between the size of step a contour may take at any one iteration and the need to control the likelihood of the contour moving out of the
neighbourhood of one local minimum of the energy functional and into an adjacent minimum.

2.2.10 Element Collapse Problem

One of the problems encountered with active contour models listed in §2.2.5 concerns the collapse of finite elements and the clustering of finite difference snaxels at points of high gradient. The problem has been noted and solutions proposed by both Amini et al. (1991) and Ranganath (1992). Amini’s dynamic programming solution allows one to specify a minimum spacing between contour nodes and Ranganath’s algorithm re-samples the implied contour after every iteration thereby enforcing uniformly spaced snaxels. The problem is caused by the gradient magnitude varying along the contour so that even when the gradient of the external energy term normal to the contour is zero everywhere along the contour, there is a non-zero gradient component tangential to the contour. By clustering the elements around the points of maximum gradient magnitude, the contour’s external energy is lower than it would be if the identical contour is re-sampled to the same number of nodes but with equal node spacing. One possible solution to this problem is to integrate the external energy with respect to the normalised path length $\tilde{l}(s)$ given by:

$$
\tilde{l}(s) = \frac{1}{L} \int_0^s \sqrt{v_n^2} ds
$$

where $L$ is the total contour length.

Note: This sets the incremental path length $dl/ds$ to be scale-independent. If the division by $L$ in (2.14) is omitted then it has been found that the contour becomes unstable and may generate ‘pseudopodia’ especially where there are gently sloping regions of the external energy function.
The normalisation suggested by Cohen (1991) and stated previously in (§2.2.8) should not suffer from the problem of collapsing elements. Karaolani et al. (1990) propose a penalty term to attempt to balance the forces causing this effect. Theoretically, the elastic energy term prefers elements to be at least equally spaced (the energy is a sum of squares of node-lengths) and the bending energy term promotes straight and equal-length elements. However, increasing the internal parameters also has the effect of causing the contour to shrink and even collapse totally.

2.2.11 Modified Internal Energy and Forces

Another disadvantage of the original active contour model given by (2.1) is that both of the internal energy terms, i.e. the tension and stiffness terms, are scale-variant so that the internal energy can always be reduced by a contraction of the contour. With reference to closed contours, Xu et al. (1994) propose that the internal force components normal to the contour that cause shrinkage be counteracted by adding a weighted internal pressure term similar to that previously used by Sullivan et al. (1990) and Cohen (1991). Sullivan reported the use of variable tension coefficients so that under the action of internal pressure the default shape of the active contour was elliptical. Cohen proposed the pressure term as an active component that was used to inflate the contour from a small trivial initialisation. A problem with the proposal of Xu is that if one requires control of stiffness in a contour, then it is necessary to allow production of bending forces normal to the boundary, which straighten or at least smooth parts of the contour. At certain points along the contour these forces will be generally oriented into the object and at other points the forces will be directed out from the object. However, the notion of using an internal varying pressure-like force to prevent collapse of a closed contour is a useful addition to the model. One may use the internal pressure term to balance only the contraction effect of the tension forces, ignoring any contraction effect of the stiffness forces. The default shape defined by a pressurised contour with non-zero internal energy is a circle. Sullivan et al. (1990) state that this makes the contour behave like a bubble where the pressure to tension ratio is proportional to the area. It is thus
possible to calculate the internal pressure required that would maintain contour equilibrium at a pre-defined area. A third method proposed by Radeva and Marti (1995) is to base the internal energy terms on the difference between a static reference contour and the active contour. The following internal energy terms replace the ones given in (2.1):

\[
E_{int}^* = \frac{\alpha}{2} \left| \mathbf{v}_s - \mathbf{v}_s^0 \right|^2 + \frac{\beta}{2} \left| \mathbf{v}_{ss} - \mathbf{v}_{ss}^0 \right| ds
\]

where \( \mathbf{v}_s^0 \) is the static reference contour.

The advantage of this method over the pressurised ‘balloon’ or ‘bubble’ model is that the contour defaults to the shape given by the reference contour and not a circle. Keeping the reference contour static means that the model will be scale-variant.

2.2.12 Region Pressure Energy

Ivins and Porril (1994) reported a region growing version of the active contour model (‘statistical snake’) that replaced the external gradient energy with a pressure energy term weighted locally along the contour by some distance metric of the external potential field at each point. Using a gray level metric the contour model could be converted into a centroid-linkage region growing algorithm with boundary smoothness constraints and this could be further generalised to respond to texture energy potentials (Laws, 1979) by using the Mahalanobis distance as a metric. If the external energy field was defined to be constant everywhere, the region energy pressure simply equalled the classical pressure energy. Thus the ‘statistical snake’ can be regarded as a generalisation of the pressurised snake. Whilst this is a powerful development of the capability of active contour models it clearly suffers from the same problems that beset standard region growing algorithms when applied to wounds: the wound region is required to have uniform gray-level, colour or textural properties and this is not generally the case.
2.2.13 Directional External Energy and Forces

The external energy term defined in (2.1) treats the potential function in an isotropic manner. It is sometimes undesirable to allow all edges to be able to attract a neighbouring part of the contour since they may represent the boundary of an object separate to that being delineated. This is the case in many wound images where strong edges are created by the contrast between the edges of a limb and the background of the image. If part of the contour is placed too close to such an edge it will be deformed to fit the edge of the limb rather than that of the wound. Several authors have formulated directional energy terms to enable the contour model to discriminate between edges based upon their orientation relative to the contour. In the case of wound images it is often found that the wound is darker than the surrounding skin and the surrounding skin lighter than the background, so that gradients of the potential function directed normally outwards from the contour are considered valid wound edges and gradients oriented contrariwise are considered spurious. Radeva and Marti (1995) use distance propagation to create the potential function, so that the potential at any point in the image is a measure of the Euclidean distance to the nearest detected edge pixel. This distance measure is given a polarity depending upon the orientation of the nearest detected edge point. They derive forces from this potential function, given by:

\[
F(v) = \text{sgn}(P(v))(\cos(\phi)P_x(v), \sin(\phi)P_y(v))
\]

where \( \phi \) is the angle of the vector normal to contour, directed into the (closed) contour and \( P_x, P_y \) are the spatial derivatives of \( P \).

This formulation provides the contour model with the ability to be repelled away from gradients with incorrect orientation, although it does not appear to be rotationally symmetric. The function may be modified to allow a degree of intransigence towards such gradients instead of the repulsion. A possible disadvantage of (2.16) is that it is not expressed in terms of energies but in terms of modified force vectors. There does not seem to be a tangible potential function whose partial derivatives yield the force vector.
function given, although this should not be taken to imply that such a function does not exist. The lack of an energy formulation reduces the number of options available for conducting the energy minimisation process; i.e. one cannot use a greedy algorithm (§2.2.9) or the Rayleigh-Ritz finite element method.

A directionally weighted potential energy function is reported by Lai and Chin (1993). Since they use a greedy algorithm, which directly minimises energy by local searches, only the energy function needs to be specified. Their potential energy function is given by the following:

\[ E_{ext}(v^i) = 1 - \frac{1}{2} \left( 1 - \cos \phi_i \right) \]

where \( \phi_i \) is the angle between the image gradient vector at snaxel \( v^i \) and the corresponding tangent vector to contour, and the magnitude of the image gradient normalised to \([0,1]\). The authors do not state whether this normalisation relates to the whole image or just to the points with which the contour is in contact at each iteration step.

2.2.14 Setting the Contour Regularisation Parameters

It was stated in §2.2.4 that the setting of the contour regularisation parameters, \( \alpha \) and \( \beta \), which control the elastic and bending energy terms of the contour respectively is problematic. Terzopoulos (1987) has suggested that the parameters be dynamically adjusted along the contour at each iteration in order to conform the contour towards a prescribed 'natural' state having a 'natural arc length', \( L(s) \) and a 'natural curvature', \( C(s) \). The formulation is given in the continuous case and the parameter (functions) are set by the following:
(2.18) \[ \alpha(s) = |v_s| - L(s) \]

(2.19) \[ \beta(s) = \kappa(s) - C(s) \]

where \( |v_s| \) is the incremental arc length of the contour and \( \kappa(s) \) is the curvature of the contour.

Leymarie and Levine (1993) state that altering the parameters at each iteration using (2.18) and (2.19) has the tendency to make the snaxels oscillate around the prescribed values of natural arc length and natural curvature. This has clear implications for the stability and convergence of the model. They propose a clipping function to constrain the values of \( \alpha \) to remain within a prescribed band of acceptable values. The curvature term poses a much more difficult problem (Leymarie and Levine, 1989). To avoid this problem they suggest setting \( \beta \) to a small positive constant, giving the contour a weak tendency to straighten.

Williams and Shah (1992) compare several discrete measures of curvature and propose the following curvature estimate as the most representative value of discrete curvature:

(2.20) \[ C(i) = 2(1 - \cos \theta) = \frac{|u^{i+1}|}{|u^{i+1}| - |u^i|}^2 \]

where \( u^i = v^i - v^{i-1} \) is a vector joining two neighbouring snaxels and \( \theta \) is the angle between two such vectors sharing a common snaxel.

They propose setting \( \alpha = 1 \) and \( \beta = 1 \), with the gradient term weighted by \( \gamma = 1.2 \) (the gradient energy is normalised locally so that it cannot exceed 1). The stiffness parameter is set to zero at points where the discrete curvature is a local maximum along the contour, provided that (a) its value is greater than a minimum curvature threshold, and
(b) the edge strength at that point is above a gradient threshold. This allows the contour to develop a corner at such a point.

Lai and Chin (1993) make use of the minimax principle (Gennert and Yuille, 1988), which they show obviates the need to explicitly set the parameters. The application of the minimax principle assumes that the energy of an active contour model in equilibrium is a continuous convex function of the regularisation parameters. It proposes that the maximum point on the surface of this function corresponds to the parameter levels that provide the best trade-off between noise-insensitivity (high regularisation) and the need to allow enough flexibility of the contour to adequately follow the desired contour (low regularisation). Applying the minimax principle, the energy term to be minimised at each iteration is the maximum of the energy terms that constitute the model.

2.2.15 Initialisation of Active Contour Models

The iterative equations for solving an active contour problem require an initial approximation to the desired solution. In this work, the initialisation relies upon a manual delineation of the boundary so that the active contour model is used to refine the solution proposed by a trained delineator.

Cohen (1991) describes a ‘balloon’ active contour model that uses an internal pressure force – which by definition acts with constant pressure normal to the boundary everywhere – to inflate a contour situated inside the boundary of the desired object. The pressure and image potential forces were controlled so that noise pixels did not trap the contour before the desired boundary was located. The need to simply initialise the contour to be a small region within the region of interest is an attractive property of this model. Unfortunately, in the context of wound measurement, the model possesses two overriding disadvantages:
Along a region of concavity of the boundary, the elastic and pressure forces will add to produce a much-increased force acting outwards from the boundary whereas at regions of convexity the elastic forces will oppose the pressure. Thus setting the elastic constant so that it is large enough to regularise the contour and also small enough to avoid causing the pressurised contour to overrun concave boundary edges could be problematic.

Many wound images have strong edges located within the site of the wound so that contour growing from within the wound site is likely to be unsuccessful in these cases. This does not preclude the initialisation of the contour outside of the wound and applying a contractile (inverse) pressure force. However, the situation is simply reversed: at concavities the pressure will cancel with the elastic force and add at convexities.

Lai and Chin (1995) show that the Generalised Hough Transform (GHT) can be used to model a contour. They use the GHT to obtain an approximate solution and state that the GHT essentially applies a rigid template model to the problem. The active contour model (deformable template) then refines the result of the GHT. This approach is most useful when the object(s) of interest are of similar shape such that they can be adequately modelled — if only approximately — by the GHT. However, in the case of wounds, the shapes are too diverse for a GHT to be applicable. However, it may be that once a wound has been delineated and this delineation refined by application of an active contour model, then the shape information gained by so doing could be employed to refine subsequent wound measurements and even to subsequently dispense with the need for delineation. If an initial wound template could be used as a reliable deformable model for subsequent measurements then the g-snake technique (Lai and Chin, 1995) could be used to provide more robust and automated wound measurement. To implement this requires a controlled study of a large enough set of wounds to ensure the broadness of its applicability.

The goal of automated initialisation could be pursued by using region-based methods such as pixel classification schemes based on colour properties of the wound.
components and surrounding skin. As previously stated, these properties vary from one patient to another so that supervised classification is not flexible enough to allow discrimination in many cases. Clustering techniques are better suited to the problem, although it is foreseeable that there will be very many instances where such initialisation will not place the entire contour in the vicinity of the wound’s boundary, since there are ‘false’ edges outside and more commonly inside the wound.

2.2.16 Global Energy Minimum Seeking

An active contour model algorithm attempts to iterate to convergence at a local energy minimum. Regularisation of the contour itself and also of the potential image function suppresses the many noisy energy minima that exist in the vicinity of each true solution. In an image with other edges beside those belonging to the object of interest, other local energy minima will exist and if the contour model is not correctly seeded, will converge to the wrong solution and commonly, to a hybrid solution with parts of various different boundaries forming the whole.

If enough information can be provided to make the energy of the true wound contour the global minimum out of all the energy minima existing in the domain of the problem, one may attempt to seek such a minimum using a trivial initialisation, e.g. a circle centred at the centre of the image of some default size, perhaps relative to the image itself. This implies that not only is the energy model both specific to wounds and general enough to encompass all wounds but also that the algorithm used to fit the model can find the global minimum. One method of finding a global minimum is that of simulated annealing (Aarts and Laarhoven, 1987). Some experimentation with this method has shown that a major factor in the likely success of this approach is an appropriate parameterisation of the contour. Storvik (1992) used simulated annealing to solve an active contour optimisation problem, but allowed only localised deformations to be made to the contour.
Staib and Duncan (1992) parameterised the boundary of an object using a Fourier decomposition which generalised to be scale and rotation invariant. They used an independent multi-dimensional Gaussian probability distribution to model the expected variation in the Fourier-decomposed parameter vector elements for their search shape. Using Bayes' theorem the boundary finding model is cast as a maximum a-posteriori problem. Taking the maximum a-priori boundary vector as an initial solution they used a gradient-descent algorithm to converge to, firstly, a coarse-scale solution achieved by blurring the image with a Gaussian filter, then using the converged result as a re-initialisation, a second finer-scale solution was sought. Use of the Gaussian probability function implies that the mean and variance of the parameters for a particular domain of possible templates or corresponding object boundaries are known.

In the case of finding a wound's boundary, the parameter distributions could be estimated from the manually delineated boundaries of a set of wounds. This process is likely to require many training samples of wound boundaries, especially if the number of parameters required to model the distributions is large. Specifying a distribution of the parameters covering all encounterable wounds is likely to yield a vague description of any one wound case.

2.3 Summary

This chapter has appraised the various segmentation algorithms that are commonly reported in the available literature. This work maintains that the varied visual properties of a substantial proportion of leg ulcer images limits the applicability of many algorithms which require the data representing each image region to be homogeneous with respect to some criterion. Indeed the principle of dividing an image into homogeneous regions is commonly cited by standard texts on image processing. Leg ulcers tend to be somewhat varied and slough formed from dead tissue complicates their appearance. In addition to this, epithelialisation tissue and the quality of the surrounding 'undamaged' skin is varied among different wounds. However, the leg ulcer images
have been gathered from several different sources and under varied illumination conditions. Improving the stability and homogeneity of illumination is certainly a worthwhile avenue of further work. The active contour models appear to be the most widely applicable technique for measuring the size of a leg ulcer. There is no assumption that the wound is homogeneous; instead the requirements for successful measurement are fairly good contrast around the wound (especially at points of high boundary curvature) and good initialisation. The following chapter formulates four active contour model algorithms that incorporate many of the proposals made by other researchers to improve, adapt and extend the capabilities of active contour models.
The preceding chapter introduced and appraised the various active contour model forms presented in the available literature. The purpose of this chapter is to describe the adaptation of the most appropriate of these models for the wound area measurement problem. Four different active contour models are developed and presented with the objective of comparing their performance as wound area measurement algorithms. The models may be divided into two categories according to the two methods of solution that they employ: (a) iterative solution of matrix equations as presented by Kass et al. (1987) and (b) direct iterative energy-minimising search using a 'greedy' algorithm (Williams and Shah, 1992; Lai and Chin, 1993). All four models require reasonable initialisation in order to produce repeatable results. Common to all four models is the construction of the external potential energy image, based on edge strength. Additionally, all algorithms make use of a common scale-descent procedure that is described in the following section. The algorithms have been implemented in the C/C++ language and run under the DOS system environment using a 32-bit protected mode interface. The programs employ VESA direct video mode addressing for graphical output.

### 3.1 Producing the External Potential Energy Field

Multi-scale gradient operators and edge detection schemes were discussed in Chapter 2, where the use of gradient operators for generating the external energy term of equation (2.1) was considered. This section further considers the use of such operators for producing the external energy function and provides some examples of the results using images of leg-ulcers. These examples are used as a discussion point to highlight some of the problems that the proposed algorithms encounter in
practice. Application of a smoothed gradient filter (e.g. derivative of Gaussian) to a gray level or intensity image produces two main side effects:

- The peak energy of an edge and associated peak attractive forces are depleted by the spreading effect of the filter. This alters the energy balance of an active contour model.
- The position of an edge may be translated by smoothing with the effect that the area enclosed by a contour varies with scale.

3.1.1 Depletion of the External Energy and Attractive Forces

The Derivative-of-Gaussian gradient operator provides regularisation by suppressing the effects of high-frequency image components that tend to be heavily noise-dominated. High frequencies responsible for image detail and fine texture (if present) are not considered to be essential requirements for accurate area measurements. Increasing the amount of regularisation by increasing the filter width parameter (scale) removes many local energy minima present in the noisy and detailed image and helps the active contour to converge to a more closely-packed set of results, i.e. increased precision. The final solution will be an optimised one, where both internal contour-regularising energies (controlling to some extent the contour shape) and external energies contribute. Therefore, it is necessary to understand the effect image regularisation has upon the magnitude of the external energy, since any overall shifts in this energy due to scale equate to a modification of the internal contour regularisation parameters. Consider therefore, the idealised 1-D step edge model shown in Figure 3.1(a) along with the profile of this edge after low-pass filtering with a Gaussian kernel at scales (i.e. standard deviation) \( \sigma = 0.25, 0.5 \) and 1.0. The derivatives of these smoothed edge profiles are plotted in Figure 3.1(b). The effect of increasing the filter scale (i.e. its width) increases the regularisation of the image and this serves to suppress the image gradient magnitude. The peak energy at the centre of the edge response is therefore depleted
with increased scale, so that the relative magnitude of the active contour's internal energy increases, thus increasing regularisation of the contour itself. One would expect to increase contour regularisation when the image (external energy field) becomes more noisy, and not less noisy as is the case when image-regularisation is increased.

It is necessary to compensate for the reduction of peak energy by scaling the gradient by the filter scale. The gradient magnitude of a unit step edge convolved with a Gaussian derivative filter of scale $\sigma$ (in one dimension only) is given by:

\[
(3.1) \quad \left( \frac{d}{dx} G_\sigma(x) \right) * H(x) = G_\sigma(x)
\]

where $G_\sigma(x)$ is a Gaussian smoothing filter.

The peak gradient in (3.1) occurs at $x=0$ and has the value:

\[
(3.2) \quad \max_x(G_\sigma(x)) = G_\sigma(x = 0) = \frac{1}{\sigma \sqrt{2\pi}}
\]
Thus compensation for peak energy variation may be achieved by multiplying the gradient result by the factor $\sigma\sqrt{2\pi}$. If the original external energy term suggested by Kass et al. (1987) is used (square of gradient magnitude) then one must multiply by the square of this factor to compensate.

Once the change in peak energy with scale is stabilised, the effect of increasing scale is simply to increase the spread of an edge's 'energy' within its locality in the image. This spreading, even though the normalised peak energy is the same at all scales, causes (a) the reduction of peak edge-attractive force due to the energy-gradient being reduced (Figure 3.4) and (b) increases the distance that the peak forces occur from the edge location (for a Gaussian profile the derivative magnitude peaks at $\pm\sigma$). For a given set of contour regularisation parameters increasing the scale increases the ability of the contour to overcome weak edges by weakening the peak attractive force of the edge. This effect is implicitly overcome by using normalisation. The external force normalisation of Cohen (1991) is suitable for this process when the minimisation algorithm is based on gradient descent. This normalisation implies that the external energy profile is piecewise linear with unit slope and with short quadratic segments forming a smooth ($C^1$ continuous) variation across gradient peaks. Algorithms that directly minimise the energy terms by iteratively searching through various control vertex states should use (3.2) to compensate for the scale-effect, unless a more dynamic normalisation scheme is used, e.g. Williams and Shah (1992). Normalisation is a key element to the successful application of the minimax principle (§2.2.14), as employed by Lai and Chin (1993).

3.1.2 Translation of Edges

Multi-scale edge detection schemes and their associated gradient operators, such as the Canny algorithm (using derivative of Gaussian filter) allow one to observe the more dominant edges at high scales free from the clutter that exists at low scales.
Blurring an image in this way has the effect of moving edges away from their original position in the unblurred image. Two modes of movement can occur when detecting edges at high scales:

- The edge contour becomes progressively smoother as scale increases.
- Edge contours can be progressively displaced with the effect that size of region defined by closed contours can change.

Clearly, the fact that contours are smoothed by blurring is unsurprising. One could utilise this property to allow the use of fewer elements to define a contour at higher scales and also to suggest that the contour regularisation parameters could be relaxed. Figure 3.2 shows two examples of leg ulcer images where the edges detected after applying a derivative-of-Gaussian gradient operator and subsequently suppressing non-directionally-maximal gradients are translated as a result of image smoothing. The darker lines correspond to the edges at higher scales. This effect must be considered by any algorithm that uses coarse-scale filtering as part of its object boundary-detection scheme.

Figure 3.2 Movement of detected wound edges with different levels of filter scale.
Chapter 3

Algorithm Development

The image used to produce the edge profiles on the left of Figure 3.2 is a well-defined wound (see 1.1 in Plate 1). The darkest contour is the edge profile that is detected with a filter scale of $\sigma=16$ pixels. The effect of image smoothing is substantial in this image. The image used to produce the edge profiles on the right of Figure 3.2 is somewhat less well-defined (see 1.4 in Plate 1). The movement of edges appears to be somewhat less affected by the application of smoothing, although there is a double contour on the left hand side of the wound which is evident at low scales. This can cause problems when manual delineation is used to initialise the contour.

3.1.3 Scale Descent Algorithm

Kass et al. (1987) and Leymarie and Levine (1993) have reported active contour model algorithms that first seek equilibrium at a high level of scale. The use of such a procedure allows the model to become less sensitive to distortion of the resulting contour caused by noise. In the case of measuring the area of a leg ulcer, there exists the additional problem of multiple local energy minima corresponding to edge segments positioned near to the boundary of the wound. These are caused by the contrast between epithelialisation tissue and blood and the surrounding skin. Such nearby edge segments that do not form part of the wound boundary are referred to as 'ambiguous'. It will be shown in Chapter 4 that ambiguities in the image also lead to manual delineations that are biased. It is considered that initialisation of an active contour model is the most difficult aspect to resolve in the case of wound measurement, therefore manual delineation is used to initialise the algorithm, implying the possibility of biased results. Starting at a high scale merges multiple edges together and overcomes the problem of varied initial contours becoming trapped by different edges. Thus contours corresponding to biased delineations are converged before seeking contour equilibrium at a lower scale. In order to be specific about the manner in which multi-scale information is combined, the scale descent algorithm is defined overleaf.
**Scale Descent Algorithm**

Initialisation: Manually delineated contour (using colour image of wound)

1. Set initial scale parameter, scale reduction factor and end scale
2. Create potential energy image using scale factor to set filter scale
3. **Iterate initial contour to equilibrium***
4. Lower scale by reduction factor

WHILE scale >= end scale
   5. Create potential energy image using scale factor to set filter scale
   6. **Iterate contour to equilibrium***
   7. Lower scale by reduction factor

* The iteration procedure used in steps (3) and (6) is one of the four algorithms described in the following sections.

The selection of the gradient filter's start and end scale parameters for each algorithm is considered in Chapter 5. The active contour model algorithms proposed to fulfil steps (3) and (6) are described in the following sections.

### 3.2 The Ground-Attract (GA) Algorithm

This algorithm combines several contributions that have been made to the active contour paradigm. Firstly, it employs the 'ground state' initial model due to Radeva and Marti (1995) and a modification of the directional-gradient external energy term due to Lai and Chin (1993). The model is spatially discretised using the finite element method, a step previously reported by Karaolani et al. (1990) and Cohen and Cohen (1993). Before defining the energy functional and its associated Euler-Lagrange equation, it is necessary to specify some component terms:
Let \( \hat{n}(s) \) be a unit vector oriented normally outwards from the contour \( v(s) \).
Let \( v^0(s) \) be a static reference contour given by the initial contour position (this defines the 'ground-state').
Let \( P(v) \) be an external energy potential function, \( P(x,y) \), sampled along the contour \( v(s) \). \( P(x,y) \) is a smoothed gradient image given by:

\[
(3.3) \quad P(x,y) = -\nabla \left( I(x,y) * G_\sigma(u,v) \right)
\]

where \( I(x,y) \) is the intensity plane of the wound image and
\( G_\sigma(u,v) \) is a Gaussian smoothing kernel with standard deviation \( \sigma \).

The energy equation and the modified Euler-Lagrange equation for the complete model, with periodic boundary conditions, are respectively given by:

\[
(3.4) \quad E_{snake}^*(v) = \frac{\alpha}{2} \int_s \left| v_x - v_x^0 \right|^2 + \frac{\beta}{2} \left| v_{ss} - v_{ss}^0 \right|^2 - \hat{n} \cdot P(v) \, ds
\]

\[
(3.5) \quad F_{snake} = -\alpha \left( v_{ss} - v_{ss}^0 \right) + \beta \left( v_{ssss} - v_{ssss}^0 \right) + \left( \hat{n} \cdot \hat{P} \right) \nabla |P|
\]

where \( \hat{P} \) is the unit-vector in the direction of gradient \( P \).

The first two terms in the integral of (3.4) are the terms proposed by Radeva and Marti (1995) which cause the contour to tend to maintain its initial size, shape and orientation. Note that the gradient term in (3.5) is not the exact variation of the energy term in (3.4), but rather an intuitive approximation. The direction of the gradient of the potential function in (3.4) is preserved whilst being gradually cosinusoidally weighted by the alignment of the normal vector to the contour and the direction of the potential gradient. Also note that it is possible to implement the minimisation of (3.4) by a direct search algorithm using both dynamic programming methods and a greedy algorithm (§2.2.9).
This active contour model may be spatially discretised by applying either the Galerkin weighted residual finite element method to (3.5) or the Rayleigh-Ritz method to (3.4) – see Stasa (1985) for a general description of these two methods. In general, the two methods can be shown to produce the same result given corresponding differential and variational formulations. However, the external term in (3.4) is not the exact equivalent of the external term in (3.5). The current implementation uses the Galerkin weighted residual method to produce the external forces. Appendix A provides the derivation of the finite element equations given the original energy functional (2.1) of Kass et al. (1987). The shape functions selected for this process are unit-spaced cubic B-Splines having the properties of (a) $C^2$-continuity and (b) at least twice differentiable. The modifications proposed in this section do not require the process of deriving the finite element matrix equations again. The assemblage stiffness matrix, $K$, for the system of equations is unchanged by the introduction of the $v^0$ term in the energy functional (3.4). However, the resultant matrix equation which is satisfied at equilibrium becomes:

$$K(V - V_0) + f(V) = 0$$

where $V_0$ is a two-column matrix of vertices controlling $v^0(s)$ and $f(V)$ is the assemblage of the elemental external force vectors given by

$$f_e(V) = \int_0^1 N^T \left( \hat{n} \hat{P}^T \right) \nabla \left| P \right| ds$$

The external forces acting at each control vertex in $V$ are integrated using the extended trapezoidal rule of closed type (Press et al., 1988). This basic quadrature formula amounts to an approximation of the external potential function, sampled along each element, by piecewise linear segments. At present, 25 equally spaced points are used. The accuracy could be improved and the number of points reduced (hence lower computational effort) by using a quadrature of higher order and Gaussian form, e.g. Gauss-Legendre, as proposed by Karaolani et al. (1992).
The equation to be solved at each iteration time-step becomes (using the semi-implicit scheme):

\begin{equation}
(I + \tau K)V_{t+1} = \tau KV_0 + V_t + \frac{f(V_t)}{|f(V_t)|}
\end{equation}

The internal forces acting at the first iteration, \( t=0 \), are zero since \( V_{t=0} = V_0 \).

### 3.2.1 Stability of the Solution

The stability of the solution may be assessed by considering the fact that the \( KV_0 \) term in (3.5) is a constant term given temporally-constant values for \( \alpha \) and \( \beta \). Thus, provided a scheme that varies these two parameters based on some merit function of the contour at each iteration is not used to iteratively tune them, the constant has no effect on the stability of the solution. The stability properties of (3.5) may therefore be considered identical to the case given in (§2.2.8). Thus the semi-implicit scheme is stable with respect to the internal parameters and may be stabilised with respect to the external forces by (a) normalisation and (b) bi-linear interpolation as used by Cohen (1991).

The acronym 'GA' stands for Ground-Attract, which expresses the two incorporated ideas due to Radeva and Marti, and Lai and Chin, of (a) ground state contour and (b) external forces that selectively attract contours based on mutual orientation, but do not repel an inversely oriented contour.

### 3.2.2 Properties of the Model

The inclusion of the \( v^0 \) term in (3.4) weighs the final solution towards the initial solution since the internal energy is dependent upon the first and second order
deviations of the active contour from its initial state and the internal parameters. The properties of this formulation are:

- It does not cause the contour to collapse in the event of over-regularisation of the contour.
- It is translation invariant, though not scale invariant.
- It biases the contour element lengths towards their initial lengths.

A possible disadvantage of the internal energy term is related to the need for the contour to deform to fit the smoothed contour of the potential function at high scales. The use of initialisation at a high image smoothing-scale is to enable the contour solution to become less dependent upon its initial state by removing many local minima in the potential function. There are two factors to consider:

- The wound boundary in many images is often ambiguous with multiple edges being produced at low scales. These edges merge and move as the scale increases.
- The manual delineation process used to initialise the contour is subject to variability (see Chapter 4). This can cause the contour to span ambiguous boundary regions.

The initial filter scale used to create the potential function must be large enough to cause neighbouring ambiguous boundary edges to merge, which implies some movement of edges. It is evident from Figure 3.2 that the apparent displacement of wound boundaries at high smoothing scales can shrink the enclosed area. Increasing the regularisation of the contour rather than the image is not a viable alternative in this case since it causes the contour to maintain more of its initial delineation error. Thus it is expected that the performance of this algorithm will be optimised at a relatively low scale. This does not imply inferiority however, since the wounds encountered in practice are so varied that some algorithms should perform better on certain types of wound. One expects the shape preserving property of the GA
algorithm to allow it to be less susceptible to catastrophic failure, i.e. collapse of the contour or for the contour to converge to a wildly different solution, far away from its initial position.

3.2.3 Minimisation Algorithm

The matrix equation derived for the Ground-Attract model is given by (3.6). This section describes the algorithm used to iterate the model to convergence using a Cholesky solution of the semi-implicit equation of (3.7). The algorithm implements a time-step reduction scheme once in the vicinity of the final solution, so that the displacements of the control vertices are refined and accuracy improved over the last few iterations. The time step parameter starts at 1 and is reduced to 1/32 in octave steps. The time step is halved when the external energy of the contour increases (Leymarie and Levine, 1993) and adds the computational burden of having to calculate the external energy term in (3.4) at each iteration to check for near convergence. The algorithm is as follows:

**Ground-Attract Convergence Algorithm**

1. Set $T=1$
2. Given the initial control vertices, $V_0$, calculate initial $E^*_{ext}$ component with last term in (3.4)

   WHILE $T>1/64$
   
   (3) Calculate new control vertex position vector using Cholesky decomposition of (3.7).
   (4) Calculate new external energy, $E^*_{ext}$, with last term of (3.4)

   WHILE $E^*_{ext}$ decreases
   
   (5) Update current control vertices using result of (3) or (6).
(6) Calculate new control vertex position vector using Cholesky decomposition of (3.7).

(7) Calculate new external energy, \( E_{\text{ext}}^* \), with last term of (3.4)

\[
\text{END while}
\]

\[
\text{Halve time step, } T.
\]

\[
\text{END while}
\]

This algorithm is implemented as a subroutine of the higher-level scale descent procedure. The scale descent algorithm is common to all four active contour models described in this chapter and is described in (§3.1).

3.2.4 Possible Improvement of the Model

The perceived disadvantage of the algorithm being unable to properly converge at high scales could possibly be overcome by allowing the static reference contour (i.e. the 'ground state') to move adaptively as the algorithm proceeds. Alternatively, one could compute the internal forces as a weighted function of the last \( m \) iterations. For instance, the ground state could be assigned the highest weight with successive moves being given less weight to guard against the external forces pulling the contour too far away from its initial position. The general balance of the weights would probably need to be set empirically given the diverse nature of leg ulcer images. It is foreseeable that the weights would need to sum to unity, otherwise the effect of adding extra terms would be to increase the internal regularisation of the contour.
3.3 The Tangent-Normal (TN) Algorithm

The *external* forces that act upon the contour in the standard active contour model can cause the elements to collapse in cases where the potential gradient is non-uniform along the boundary (cf. §2.2.10). Additionally, the *internal* elastic energy term in the standard model is responsible for producing scale varying forces that can cause the whole contour to collapse to a point (§2.2.11). This algorithm proposes two modifications to the standard active contour model to overcome some of the associated problems, although the modification to the elastic term will not remove the scale dependence.

- *Xu et al.* (1994) have proposed use of an internal pressure force acting normally to the contour and weighted at each point such that the components of the elastic and bending forces at these points are zero. The algorithm proposed in this section makes use of such a notion to completely cancel out the normal component terms of the *elastic* forces. This does not, of course, nullify the component of this force that acts tangentially to the contour. Following the discussion in §2.2.11, the pressure term proposed here is not formulated to oppose the *bending* forces, which remain identical to the original definition (see §2.2.3.2).

- The second modification made is to consider only the component of external forces that acts normally to the contour. Thus, instead of cosinusoidally weighting the external forces (§2.2.13), the proposal made here is to completely remove the tangential component of the standard external force that causes contour elements to bunch up at points of locally maximum gradient.

The algorithm is referred to as the ‘TN’ (Tangent-Normal) algorithm due to the combination of tangential internal elastic forces and normal external forces. Considering these two modifications, the Euler-Lagrange equation for the model may be written as:
\[ F_{\text{snake}} = \alpha \mathbf{v}_{ss}^\parallel + \beta \mathbf{v}_{ssss} + (\nabla P(\mathbf{v}))^\perp \]

where \( \mathbf{v}_{ss}^\parallel \) is the component of elastic force \( \mathbf{v}_{ss} \) tangential to \( \mathbf{v}(s) \) and

\( (\nabla P(\mathbf{v}))^\perp \) is the component of external force normal to \( \mathbf{v}(s) \).

Note that since (3.8) is formulated in terms of forces, an equivalent energy formulation is not specified. This prevents the equation from being implemented by both direct minimisation algorithms and the Rayleigh-Ritz finite element method. Two remaining possible options are finite differencing and the Galerkin weighted residual method.

### 3.3.1 External Force Made Normal to the Contour

In the standard active contour model the external force components resulting from the potential function are simple spatial derivatives of the function, thus the external forces, \( \nabla P(\mathbf{v}) \), always act in the direction of maximum gradient of the potential function, irrespective of the relative orientation of the contour. The external force vector, derived from the potential function, is redefined here so that it is always oriented in a direction normal to the contour. This is achieved by simply taking the component of the potential gradient in the direction normal to the contour. Let \( \mathbf{n}(s) \) be a unit normal vector to the contour \( \mathbf{v}(s) \). Then the magnitude of the gradient potential in the direction of \( \mathbf{n}(s) \) is given by:

\[ (3.9) \quad (\nabla P(\mathbf{v}))^\perp = \nabla P(\mathbf{v})[\mathbf{n}^T \mathbf{n}] \]

where \( \mathbf{n}^T \mathbf{n} \) is the outer-product of the unit normal vector with itself.

Note that \( P(\mathbf{v}) \) is a scalar quantity, and thus (3.9) does not consider the orientation of the gradient of the gray level image used to produce the potential. This may be
compared to the external force specification for the GA algorithm (3.5): In (3.9) the cosinusoidally-weighted forces act normally to the boundary, whereas in (3.5) the cosinusoidal weight is calculated on the basis of the alignment of the potential function (a vector term) and the contour. The forces act in the direction of the maximum gradient of the magnitude of the potential function. In (3.9) the directional information comes from the gradient of potential function. Using the Galerkin weighted-residual method (3.9) becomes:

\[
(3.10) \quad f(V_e) = \int_0^1 N^T \nabla P(V_e) \left[ \hat{n}^T \hat{n} \right] ds
\]

This equation may be solved numerically for each element with a suitable quadrature scheme. Subsequently, the force vectors corresponding to each element are assembled to produce the overall force vector.

### 3.3.2 Elastic Force Made Tangential to the Contour

The elastic forces acting along the contour may be decomposed into any arbitrary pair of orthogonal component vectors. Thus the elastic force acting at any point on the contour may be expressed in terms of a component normal to the contour and a component tangential to it. The elastic force component acting normally to the contour is that component responsible for causing the entire contour to tend to collapse to a point in the absence of competing forces (at points of concavity the elastic forces cause the contour to expand, see Figure 2.1). The force component acting tangentially to the contour is responsible for promoting equal length elements. Removal of the normal force component enables the contour to maintain its size and shape, rather than collapsing in the absence of external forces.

It should be noted that for stable non-contractile behaviour under the elastic energy model the energy term must be **scale invariant**. Scale invariance implies that the
elastic energy remains unchanged after the contour or contour element has undergone a linear magnifying transformation. Any tendency for the energy to change with scale implies the production of an elastic force vector component which will tend to drive the contour towards that scale at which elastic energy is minimised. As already mentioned, this is the case with the classical elastic energy definition which tends to drive the contour to complete collapse.

3.3.3 Implementation of Tangential-Only Elastic Force

It was shown in Figure 2.4 that the elastic forces acting along the contour have a large normal component responsible for causing the contour to collapse in the absence of large enough competing forces. The standard elastic force is defined by (2.4) and may be decomposed into any two arbitrary orthogonal vectors. Remodelling (2.4) in terms of a vector-pair formed from the component of elastic force (a) normal to the curve and (b) tangential to the contour yields the following equations that are no longer independent in terms of the $x$ and $y$ components of the contour:

\begin{align*}
(3.11) \quad F_{\text{elastic}}^n &= \left( \frac{-y_s(x_s y_{ss} - x_{ss} y_s)}{x_s^2 + y_s^2}, \frac{x_s(x_s y_{ss} - x_{ss} y_s)}{x_s^2 + y_s^2} \right) = v_{ss}[\hat{n}^T \hat{n}] \\
(3.12) \quad F_{\text{elastic}}^t &= \left( \frac{x_s(x_s y_{ss} + y_{ss} y_s)}{x_s^2 + y_s^2}, \frac{y_s(x_s y_{ss} + y_{ss} y_s)}{x_s^2 + y_s^2} \right) = v_{ss}[\hat{t}^T \hat{t}]
\end{align*}

where $\hat{n}$ is a unit normal vector to the contour and $\hat{t}$ is a unit tangent vector.

It is easy to verify that the sum of (3.11) and (3.12) equals the original elastic force equation given by (2.4). Since this algorithm requires only the tangential component (3.12), one might consider attempting to find the Galerkin weighted residual formulation to obtain the associated stiffness matrix components. Performing this
task analytically appears to be rather unwieldy and the forces would clearly be a non-linear function of the control vertices. Two alternative solutions are possible:

- Calculate the weighted residuals of the normal vectors given by (3.11) and subtract them from the external force vector. This is equivalent to specifying an internal pressure term where the pressure coefficient varies around the contour with a magnitude given by the magnitude of (3.11). The pressure is defined to act normally to the contour and also must maintain the sense of the elastic force component, i.e. into or out of the contour.

- An approximate solution may be formed by considering the components of the stiffness matrix equations directly. A set of such equations defines the lumped forces acting upon each control vertex of the piece-wise contour v(s) that are responsible for its deformation. Figure 3.3 shows such a set of forces (dashed lines) along with the weighted residuals of the tangential and normal components of the forces. The dashed lines are simply the vectors arising from calculating $K\mathbf{v}$, with $\beta=0$. The orientations of the vectors corresponding to the weighted normal and tangential components are approximately aligned with the respective normals and tangents to the contour at the element end points. Thus the normal components in Figure 3.3 are an approximation to the integrated weighted residual of (3.11) and the tangential components approximately equal to the integrated weighted residual of (3.12).

The current implementation makes use of the approximate solution. If the normal components of the vectors shown in Figure 3.3 are approximated (by calculating $K\mathbf{v}$, $\beta=0$) and taking the components normal to the contour at the element end points and subtracted from the external force vector then an explicit step approximation is obtained. This may have implications for the stability of the system since explicit components can generate unacceptably large displacements. Instead, an implicit method of approximating the normal components is used. The
implicit approximation does not allow for the displacements of the control vertices to be expressed as a sum of independent forces arising from the different energy terms (unlike a wholly explicit method). This leads to a two-stage iteration process that involves performing the standard semi-implicit iteration without removing any elastic force components, followed by a 'correction' step. The correction step contains the implicit approximation and attempts to remove the elastic normal component present in the displacements of the control vertices. Stage one solves the following iteration equation by Cholesky decomposition, where $K$ is the standard stiffness matrix:

\[(3.13) \ (I + rK)V_{t+1} = V_t + rf^\perp(V_t)\]

Stage two is defined by the following sequence of operations:

(a) Solve $\left(I + rK\right)V_{t+1}^{\text{elastic}} = V_t$
(b) Calculate displacements $\Delta V = V_{t+1}^{\text{elastic}} - V_t$
(c) Calculate the displacement components normal to the contour at the respective element end points with:

$$f^\perp_{\text{elastic}} = \text{diag}(\text{diag}(\Delta V N^T)) \hat{N}$$

where $\hat{N}$ is a matrix of contour unit normals at the element end points. The operation $\text{diag(diag}(M))$ simply removes the upper and lower triangular parts of any matrix $M$ (MATLAB® help menu).
Finally, update the control vertices with \( \mathbf{V}_{t+1} = \mathbf{V}_{t+1} - f_{\text{elastic}} \).

### 3.3.4 Stability of the Solution

The stability of the solution is more difficult to assess than the previous model because of the nature of the non-linear correction step. The displacements of the control vertices caused by the correction, however, can be no greater in magnitude than the displacements caused by the standard elastic term. Use of the implicit method to calculate the standard displacements (in the absence of any other forces) guarantees stability of the standard term. The fact that the modified displacements must be smaller than the standard displacements does not prove that the algorithm will converge, however, it does provide some notion that the model should be stable.

### 3.3.5 Minimisation Algorithm

This algorithm is similar to the Ground-Attract minimisation algorithm described in this chapter. A necessary alteration required is an extra step to remove the normal component of the elastic forces as previously described. In order to check for near-convergence and thus refine the step size parameter, the standard external energy is calculated by (3.15).

\[
E_{\text{ext}}^* = \sum_{e=1}^{N} \int_0^1 P(s; \mathbf{V}_e) ds
\]

When this energy rises, the contour should be in the vicinity of the final solution.
**Tangent-Normal Convergence Algorithm**

(1) Set $T=1$

(2) Given the initial control vertices, $V_0$, calculate initial $E'_{ext}$ with (3.15)

WHILE $T>1/64$

(3) Calculate new control vertex position vector using Cholesky decomposition of (3.13) followed by the correction step.

(4) Calculate new external energy, $E^*_{ext}$, with (3.15)

WHILE $E^*_{ext}$ decreases

(5) Update current control vertices using result of (3) or (6).

(6) Calculate new control vertex position vector using Cholesky decomposition of (3.13) followed by the correction step.

(7) Calculate new external energy, $E^*_{ext}$, with (3.15)

END while

Halve time step, $T$.

END while

### 3.4 Normalisation of Parameters

The total energy of an active contour model in some arbitrary state and with some arbitrary external energy conditions is dependent not only upon the contour regularisation parameters but also upon two other parameters introduced by the discretisation and by the smoothing of the external potential function:

- The number of discrete sampling points or the number of elements in the contour.
- The scale of smoothing filter used to create the external energy field.
Ivins and Porril (1994) presented a suitable modification of the internal energy terms to maintain the energy balance of the active contour model in the event of changing the number of snaxels or elements in the contour:

\begin{align}
\alpha &= N\alpha_0 \\
\beta &= N^3\beta_0
\end{align}

Also, since the external energy in the current implementation is integrated over each element in the range \( s=[0..1] \), the external energy requires division by \( N \) to maintain energy balance.

In equations (3.16) and (3.17), \( N \) is the number of snaxels/elements and \( \alpha_0 \) and \( \beta_0 \) are independent of \( N \). The modification is necessary because the elements are based on unit-spacing in the independent parameter \( s \). Altering the number of elements over which the contour is defined alters the magnitude of the contour's derivatives by altering the effective range of \( s \), i.e. \( s=[1..N] \). In practice, one effect of this modification is the tendency to require \( \beta_0 << \alpha_0 \) when a reasonable number of snaxels or elements are used. Currently, the number of elements is fixed at 32, which appears to adequately represent the boundaries of the wound images encountered during this work. Erring on the side of caution it is probably better to have more elements than necessary.

### 3.5 The Minimax Algorithms (MX and MG)

The previous two algorithms, the GA and TN finite-element implementations, have parameters that must be properly set in order for them to work effectively without over-regularising the contour. Making use of the minimax principle (Gennert and Yuille, 1988), Lai and Chin (1993) have shown how an active contour model may be implemented without regard to the explicit setting of what they term the
'nuisance' parameters, i.e. the tension and stiffness constants. Their implementation allows the minimax optimisation to apply locally across the contour so that regularisation is adaptive depending upon the conditions existing at any part of the contour, e.g. image noise and curvature of the contour being followed. The energy minimisation algorithm is implemented as a direct local search using a modification of the greedy algorithm (Williams and Shah, 1992) and is based upon a finite difference discretisation; i.e. the snaxels are single points rather than piecewise continuous curves. It is evident that the algorithm may alternatively be implemented using cubic elements to define the contour and thus base the energy equations on a piecewise-continuous model. This requires modification of the energy term expressions since Lai and Chin specify them in discrete form only. The following sections describe the implementation of the continuity (elastic), curvature (bending) and external energy terms using a piece-wise continuous contour definition.

3.5.1 Continuity Term

This term is related to the elastic energy term of Kass et al. (1987) although its particular formulation promotes equality of spacing between snaxels. Lai and Chin's continuity energy term (3.18) is scale-invariant; hence it does not cause the contour to collapse.

\[ E_{\text{cont}}(i) = \left( d_\infty - \| v^i - v^{i-1} \|_\infty \right)^2 / d_\infty^2 \]

where \( d_\infty \) is the averaged inter-snaxel distance under the \( \| \|_\infty \) vector norm and \( v^i \) is a contour point (snaxel).

An equivalent formulation of (3.18) in the continuous (piecewise) domain is given by:
(3.19) \[ E_{\text{cont}}(s) = \left( \frac{L - |v^e_s|}{L^2} \right) \]

where \[ L = \frac{1}{N} \sum_{e=1}^{N} \int_{0}^{1} |v^e_s| ds \] is the average element length.

Following Ivins and Porrir (1994), the number of elements, \( N \), used to discretise the contour is included in (3.19) to preserve the energy balance of the active contour model if and when the number of elements is altered.

### 3.5.2 Curvature Term

Lai and Chin measure curvature at a contour point by calculating the angle formed between two vectors subtended from that point to the two neighbouring points. Their curvature term is given by:

(3.20) \[ E_{\text{curv}}(i) = \frac{1}{2} \left( 1 - \cos(\psi_i) \right) = \frac{1}{4} \left( \frac{\|u^i\| - \|u^{i-1}\|}{\|u^i\|} \right)^2 \]

where \( u^i = v^{i+1} - v^i \) is a vector joining two neighbouring points.

The curvature energy is 0.5 when the two vectors form a right angle and approaches a maximum of 1.0 as the angle becomes increasingly acute. This is a scale invariant measure of curvature.

Curvature at a point on a continuous function can be defined as the inverse of the radius of a circular arc that is tangential to the function at that point (see Stroud, 1987). A continuous representation of curvature for a parametric contour element is given by:
(3.21) $C(s) = \frac{x_s y_s' - y_s x_s'}{(x_s^2 + y_s^2)^{3/2}} = \frac{n \cdot v_{ss}}{|v_s|^3} = \frac{\hat{n} \cdot v_{ss}}{|v_s|^2}$

where $\hat{n}$ is a unit-normal vector to the contour.

Thus curvature is equal to the magnitude of the normal component of the term $v_{ss} = (x_{ss}, y_{ss})$ divided by the square of incremental arc length. An interesting point is that when the parameter $s$ represents the true contour arc length, then the boundary tangent and normal vectors ($v_s$ and $n$ respectively) have unit magnitude, and $v_{ss}$ is also normal to the boundary. Curvature is then equal to the magnitude of $v_{ss}$. Clearly, since $C(s)$ is a function of radius, the measure is not scale-invariant and attempting to minimise an energy term that integrates (3.21) will cause the contour to expand, since minimising curvature equates to maximising the radius of curvature. A scale-invariant measure of curvature may be defined by multiplying the curvature by the incremental contour length:

(3.22) $E_{\text{curv}}^E(s) = \frac{|C(s)||v_s'|}{\pi}$

The division by $\pi$ causes (3.22) to integrate to 0.5 for a perfectly circular 90° arc (the integral is taken over the range $s = [0..1]$). The total sum energy for the entire contour is obtained by integrating (3.22) over each element in the contour and summing the resulting integrals. Integrating (3.22) with respect to $s$ is equivalent to integrating $|C(s)|$ with respect to incremental contour length. This means that a magnification of an element affects both the curvature term and the incremental length term with the net result being that the curvature energy is dependent upon shape only. Taking the magnitude of $C(s)$ is necessary to avoid negative curvatures cancelling-out positive curvatures.
3.5.3 Gradient External Energy Term

The external energy term of Lai and Chin is defined by the following equation:

\[ E_{\text{gradient}} = 1 - |\mathbf{P}(v)| \left( 1 - |\sin(\phi)| \right) \]

where \( \phi \) is the angle between the gradient vector, \( \mathbf{P}(v) \), and the normal to the contour.

This produces a much more directional gradient weighting term than the standard cosine weight (dot product) used in equation (3.4). Figure 3.4 shows a comparison of the two energy terms as functions of the angle \( \phi \). With regard to the gradient weighting proposal of Radeva and Marti (1995), it would be a simple matter to remove the magnitude operator and allow the energy function to repel inversely oriented edges. The gradient magnitude in (3.23) is normalised in the range [0..1], although it is not clear as to the domain over which the normalisation takes place. Williams and Shah (1992) normalise the gradient magnitude locally by considering, for each point, the gradient magnitudes of the 3x3 pixel neighbourhood within which each snaxel can move at each iteration. This tends to produce a strong tendency for a snaxel to remain at a point of locally maximum gradient, even if all gradients within the neighbourhood are relatively weak. They apply a minimum range constraint to the normalisation to counteract this effect.

![Figure 3.4 Comparison of directional gradient weighting](image-url)
The normalisation model proposed in this section is to scale data within the range of gradient values found along the whole contour at each iteration. Equation (3.24) expresses the proposed normalisation:

\[ P(v) = \left( |P(v)| - P_{\text{min}} \right) / \left( P_{\text{max}} - P_{\text{min}} \right) \]

where \( P_{\text{min}} = \min_s (|P(v)|) \) and \( P_{\text{max}} = \max_s (|P(v)|) \)

Combining this normalisation with a cosine-weighted gradient magnitude operator produces the definition of external energy used with this model. The energy is normalised within the range \([0..1]\).

\[ E_{\text{gradient}}(v) = 1 - |P(v)| \cos \phi \]

where \( \cos \phi = P \cdot \hat{n} / |P| \)

### 3.5.4 Gray Level External Energy Term

Early experiments have shown that ambiguous contours as detected by a gradient operator are likely to cause the result to be dependent upon initialisation. The use of a scale descent algorithm can help to make contours starting from different initialisations converge to the same result. However, this can sometimes result in a hybrid 'solution' combining sections of two different contours. In cases where the gray level is more consistent around the true boundary, an adaptive gray level seeking energy term may be used to add weight to the contour which has most consistent gray level:
(3.26) \[ E_{\text{gray}}(v) = \frac{|I(v) - \bar{I}(v)|}{n\sigma} \]

where \( n \) is an arbitrary constant, 

\[ \bar{I}(v) = \frac{1}{N} \sum_{e=0}^{N-1} I(v)ds \]

and \( I(v) \) is a gray level image function sampled along the contour \( v \).

Clearly, the comments concerning homogeneity of region information in §2.1 is relevant to this proposal. A modification which could be employed, but which has not been extensively tested is to allow the 'reference' gray level -- the mean gray level in the case of (3.26) -- to vary around the wound in a controlled fashion. Wounds which are on highly curved parts of the lower leg can exhibit a marked change in intensity due to somewhat directional illumination. An example of a suitable reference function is to sum the average gray level plus first two harmonics of gray level. This allows for a maximum of two local minima and two local maxima of intensity along the contour.

3.5.5 Continuous Domain Minimisation

The continuous domain application of the minimax principle to energy minimisation yields (3.27) for the 3-term minimax algorithm (MX) and (3.28) for the 4-term algorithm incorporating the gray level term (MG), which must be minimised among a set of local search patterns at each iteration:

\[
(3.27) \quad E_{\text{snake}}^* = \sum_{e=1}^{N} \max \left( E_{\text{cont}}(v^e), E_{\text{curv}}(v^e), E_{\text{gradient}}(v^e), E_{\text{gray}}(v^e) \right) ds
\]

\[
(3.28) \quad E_{\text{snake}}^* = \sum_{e=1}^{N} \int \max \left( E_{\text{cont}}(v^e), E_{\text{curv}}(v^e), E_{\text{gradient}}(v^e), E_{\text{gray}}(v^e) \right) ds
\]
The integration of the continuous maximum energy profile in (3.27) and (3.28) is
implemented by numerical quadrature. At present the closed-form extended
trapezoidal rule is used. In addition to the use of quadrature to perform integration
of external energy along the contour, one could also follow the example of Cohen (1991)
and use a 2-D image interpolation scheme to allow the external energy to vary
smoothly across pixels, although this has not been implemented in the current
algorithm. This process of defining contour energy by continuous domain maxi­
misation of functions is depicted in Figure 3.5 for one element of a complete contour. The area under the
solid line representing the function $\max(E_{\text{cont}}(v), E_{\text{curv}}(v), E_{\text{gradient}}(v))$ is given by
the integral term in (3.27) and this quantity is summed for all elements giving the
total contour energy. At each iteration, the greedy algorithm attempts to minimise
(3.27) among the set of local contour deformations. The modified greedy algorithm
of Lai and Chin is guaranteed to converge to a unique minimum when one snaxel
(control vertex in the piecewise continuous case) is moved at a time. After moving a
single vertex, all affected parameters, e.g. contour length, must be recalculated to
avoid the production of limit-cycle oscillations.

3.5.6 Iteration and Convergence of Direct Minimisation Algorithms

The MX and MG algorithms use the modified greedy algorithm applied to equations
(3.27) and (3.28) respectively. The modified greedy algorithm for the case of cubic
B-Spline contour elements solves the following equation for each control vertex defining the contour:

\[
E^*_e = \min_{V^e} (E_{e-3} + E_{e-2} + E_{e-1} + E_e)
\]

where

\[
E_e = \int \max_s (E_{\text{cont}}(v^e), E_{\text{curv}}(v^e), E_{\text{gradient}}(v^e)) ds
\]

and \(v^e = N(s) [v^{e-3} \ v^{e-2} \ v^{e-1} \ v^e]^T\)

This summation is taken because altering any single control vertex controlling \(v(s)\) will affect the associated four neighbouring contour elements and thus their energies, when \(C^2\) cubic B-Splines are used as the basis functions (\(N(s)\)).

The discrete point models of Williams and Shah (1992) and Lai and Chin (1993) constrain the snaxels to lie on a regular grid so that the position of each snaxel corresponds to a particular image pixel. In the MX and MG algorithms described in this section, the control vertices are free to move anywhere in the continuous spatial domain. They are moved by single pixel steps initially, since the contour is placed fairly close to the boundary by the delineator and moreover the time taken for the algorithm to execute is not at present considered paramount. However, constraining the steps to be equal to the pixel size can produce significant variances in the areas enclosed by a contour. Clearly, the effect will be scale-dependent with smaller and more-compact contours being affected to a greater degree. Thus the algorithms use the same type of convergence algorithm as employed by the GA and TN algorithms. The algorithm is described here, where \(\Delta\) is the spatial step size that is adjusted as the algorithm converges:
**MX/MG Direct Energy-Minimisation Algorithm**

1. Set $\Delta = 1$
2. Label initial control vertex vector $V_0$ as $V_{old}$

   WHILE $\Delta > 1/64$
   
   (3) FOR $m = 1$ to $N$
   
   Calculate new control vertex position vector, $V_{new}$, using (3.29)
   
   WHILE $V_{new} \neq V_{old}$
   
   (4) $V_{old} = V_{new}$
   
   (3) FOR $m = 1$ to $N$
   
   Calculate new control vertex position vector, $V_{new}$, using (3.29)
   
   END while

   Halve step-size, $\Delta$.

   END while

The combination of the minimisation equation (3.29) and the gradient normalisation method proposed in §3.5.3 appears, in practice, to contribute to the element-shrinking problem. This occurs because the difference in energy between points of minimum and maximum gradient is always equal to 1 so that points of maximum external energy always have the highest possible value of 1. This means that the 'max' function in (3.29) can reduce the external energy by selecting displacements of the vertices controlling the local element that extend the ends of the element along the contour in order to incorporate points of stronger gradient (lower energy). The regularising terms in the contour do not oppose this action because the 'max' function effectively disables them until the gradient energy is sufficiently low.
3.6 Summary

Clearly, the four active contour models described in the foregoing sections have different properties, which will affect their performance when applied to the task of leg ulcer area measurement. The following statements summarise the perceived advantages and disadvantages of the algorithms:

**GA Algorithm**

- The explicit inclusion of the initial contour in the energy equation should enable this algorithm to be robust against noisy edges at low levels of image smoothing and, at high levels of image smoothing, should guard against the shift in the wound edge if it merges with a strong nearby edge, e.g. the edge of the limb.
- However, the inclusion of the initial contour has two disadvantages: (a) it does not readily allow for a weak initialisation, where the contour is placed near to, but not explicitly on the wound contour, and (b) it is not formulated to account for the movement of edges at high levels of image smoothing.

**TN Algorithm**

- This model makes no assumptions about the shape of the model and therefore it can deform as necessary at high image smoothing scales where the wound edges will be displaced. It should not suffer from element collapse, since the external force component tangential to the contour is absent. Also, the contour should not tend to shrink under high levels of the internal tension parameter, $\alpha$.
- There is no safeguard in this algorithm to prevent against the contour following the wrong contour as the image is de-blurred: edges merged into one apparent edge by smoothing bifurcate during scale descent, and it
is thus possible that the contour may begin to follow a non-wound edge segment.

**MX Algorithm**

- The two major advantages of this algorithm (and the MG algorithm) are (a) there is no need to set parameters, thus it should naturally adapt from image to image; the parameters also implicitly vary along the contour so that weak edge segments are not subject to the same level of regularisation as strong edge segments, and (b) the energy terms are scale invariant, thus it should be able to naturally adapt to any size of wound. In addition, the model should allow for weak initialisation.
- A possible disadvantage is that the model may suffer from the element collapse problem, despite weighting the external energy according to the alignment between the image gradient vector and the contour normal. This is a weaker condition than the normal-only edge force of the TN algorithm.

**MG Algorithm**

- This model shares the same properties as the MX algorithm. However, the inclusion of the gray level energy term should enable it to better discriminate between wound and non-wound edge segments. This does assume that the gray level profile around a leg ulcer is more uniform than a contour formed partly by the wound boundary and partly by nearby 'false' edges, particularly at low levels of image smoothing.
This chapter describes two experiments undertaken to estimate the level of accuracy obtainable when wound area measurements are made by manually delineating the boundary of a wound image using a standard PC mouse. Error in area measurement obtained by wound image delineation is tentatively attributed to the effects of subjective appraisal combined with manual dexterity and the mechanical quality of the mouse. This gives rise to the possibility of differing opinions upon the location of the boundary and thus produces bias between the area measurements of different delineators. Error due to manual dexterity and mechanical imperfections of the equipment is expected to be manifest as random measurement noise. The investigation of such effects is embodied in two experiments. The first experiment aims to measure the delineation accuracy (as bias and precision of area measurements) of several volunteer delineators over some widely differing wound images. The second experiment aims to quantify the effect of display size of a wound upon delineation performance.

4.1 General Assessment of Manual Delineation Performance

4.1.1 Objectives

(a) **Delineation Bias** - To determine whether different delineators measure the same wound area for a given wound. The question posed is whether or not the delineators are, on average, in agreement with each other. If no bias is detectable then the delineators are said to be in agreement.

(b) **Delineator Precision** - To estimate the variation or range of variation of manual wound area measurements in order to gauge expected levels of repeatability.
Notes

- If a bias exists, then the accuracy of wound measurements is only completely specified if both bias and precision are specified.
- In the event of statistically significant differences of a practical magnitude being found, it makes sense to attempt to determine what image features are responsible for affecting the judgment of the sample of delineators.

4.1.2 Experimental Set-up and Procedure

A set of ten wound images was specifically selected from an existing wound-image library for the experiment, which were subjectively classified into three groups. Three of the images possessed subjectively good wound boundary contrast and relatively uniform intensity in and around the wound site. Four images were more typical of the wound images in the library in that the contrast at wound boundary was generally lower and more varied that of the first three images, the surrounding skin included granulation tissue, some slough was present and the scene illumination was evidently non-uniform. Finally, the last three images were poorly defined with little discernible boundary contrast and poor illumination. Five subjects with experience in examining wound images were asked to use a graphical computer program which allowed them to delineate the wound boundary in each image under control of a standard mouse. Each delineator was asked to make a total of five delineation attempts at each of the ten wounds. The program was written in the 'C' programming language using Borland's 'Object Windows' Library for developing applications to run under the Microsoft Windows™ operating system. The software was capable of displaying colour images of wounds in a 'true colour' graphics mode (approx. 16 million simultaneous colours). The result of each delineation attempt was an image of wound boundary pixels forming a closed border. The area of the wound was then measured by counting the number of pixels lying inside or on the boundary of the wound.
Two effects were identified in advance which could affect the results:

(a) It is considered that delineators are more likely to remember which parts of the image were previously included in the wound definition if rapidly presented with the same image several times in succession. To reduce this ‘learning’ effect, the delineators were asked to make only one attempt at each wound per day, so that the trial was spread over five days.

(b) Potentially, the delineators may lose concentration if the task of delineation takes too long, so the boundaries drawn last on a particular day may tend to exhibit more variation than those drawn first each day. The countermeasure for this effect was to randomise the order of image presentation on each day to spread any actual effect over all the wounds.

4.1.3 Hypothesis Test for Bias: Analysis of Variance Fixed-Effects Model

Analysis of Variance (ANOVA) is a standard statistical method for testing the equality of several population means given some basic assumptions about the nature of the data:

- The samples are taken randomly and the data are mutually independent.
- The populations from which the samples are drawn are Gaussian-distributed.
- The populations from which the samples are drawn have equal variances.

It has been shown by Box (1954) that for fixed effects ANOVA, the $F$-test for differences among means is "most robust for $\alpha$", where $\alpha$ is the significance level of the test. When the sample sizes are equal, Neter et al. (1996) state that the effect of unequal variances is to slightly raise the significance level of the test. They also state that the $F$-test is robust against non-normality of the residuals provided that the departure from normality is not too serious, with kurtosis having a greater effect than skewness upon the significance level. However, confidence intervals for single comparisons between
factor-level means can be greatly affected by non-normality. The most serious infringement of the assumptions occurs when the residuals are not independent of each other. Neter et al. (1996) suggest that "randomisation when serial correlations are expected can be a good insurance policy."

Outliers affect estimates of both factor-level means and, more strongly, variances. In the present case, if one of the five samples is affected by a single outlier, this can exaggerate the variance for that sample and lead to a conclusion of differing variances. There are many types of residual plots that can be used to visually detect the presence of outliers e.g. dot-plots, stem-and-leaf plots and normal-probability plots. Complementary to the subjective analysis of such plots is the Bonferroni test for outliers using studentised deleted residuals (Neter et al., 1996). This test is useful in cases where graphical methods to identify outliers are considered inconclusive.

Considering the foregoing discussion, the proposed pre-ANOVA analysis to test for serious violation of the assumptions, stated above, is as follows:

- Detection of outliers and checking of normality using plots of residuals against expected normal probabilities for each wound case.
- Hypothesis test to complement the subjective appraisal of the above plots:
  - Bonferroni test for outliers using studentised-deleted residuals.
- Use of the Burr-Foster ‘Q’ test for equality of variances. This test will determine if the experimental delineators produce equal amounts of variability, which is a partial answer to the experimental aim stated in §4.1.1(b).

4.1.3.1 The ANOVA Bias-Test Model

In the case of each of the ten individual wound images, the null hypothesis states that all delineators measure the same wound area. Specifically, the null hypothesis, $H_0$, and its alternative, $H_1$, may be stated as:
(4.1.1) \[ H_0 : \text{All } A_i \text{ are equal for } i=1..5 \text{ (delineators)} \]
\[ H_i : \text{Not all } A_j \text{ are equal} \]

where \( A_i \) is the 'true' wound area measured by delineator \( i \).

For any particular wound image, let \( a_y \) denote the \( j \)th delineated wound area observation of delineator \( i \). The test for \( H_0 \) is thus conducted using the \( a_{ij} \) data in the standard manner for an ANOVA fixed-effects model.

### 4.1.3.2 Pre-ANOVA Checks for Violation of Assumptions

(a) *Test for Outliers and Normality of Residuals*

For each wound, all \( N=25 \) delineated area measurements \( (a_{ij}) \) are pooled to produce a residual value versus expected normal probability plot with a reasonable number of data points. The residuals are calculated with (4.1.2) and the expected normal probabilities are estimated using (4.1.4). The expected normal probabilities should be plotted on normal-graph paper. The probabilities may be transformed for plotting on a linear axis scale by calculating the standardised 'z' value, \( z(P_k) \), of the expected probability (Wadsworth, 1990). The pooled plot is constructed by first ordering the \( e_{ij} \) over the entire set of residuals according to magnitude (4.1.3) and subsequently plotting \( e_k \) against either \( P_k \) or \( z(P_k) \).

(4.1.2) \[ e_{ij} = a_{ij} - \bar{a}_i \]

(4.1.3) \[ e_k = e_{ij} \text{ where } k = \text{Rank}(e_{ij}), k = 1..N. \]

(4.1.4) \[ P_k = (k - 0.5)/N \]
(b) *Inequality of Variances*

Several hypothesis tests for checking the equality of variances have been published. Bartlett’s test (Montgomery, 1991), Hartley’s test, the Levene test and the Modified Levene test (Neter *et al.*, 1996), Bartlett and Kendall’s ‘log $s^2$’ test and the Burr-Foster ‘Q’ test (Anderson and McLean, 1974). The first three tests are considered sensitive to deviations from normality whereas the modified Levene test is considered robust. The ‘log $s^2$’ test requires at least two samples for each factor level, although a sample containing a minimum of 10 observations could be randomly split into two samples.

The Burr-Foster ‘Q’ statistic for comparing the variances of $N$ samples is considered robust to non-normality of the error terms. When the samples are of equal size, the test statistic, $Q$, is defined as:

$$Q = \left( \frac{\sum_{i=1}^{i=N} s_i^4}{\left( \sum_{i=1}^{i=N} s_i^2 \right)^2} \right)$$

where $s_i^2 = i^{th}$ sample variance estimate.

### 4.1.3.3 Inspection of Images for Regions of Ambiguity: Median Boundary Method

In this experiment, where biases are deemed to exist between delineators the ‘Median Boundary’ test described here is a robust test designed to analyse which regions of each image are ambiguously interpreted when $H_0$ is rejected (4.1.1). In total, each wound image will be delineated 25 times, possibly with different regions of the wound site included or excluded from the definition of the wound. The test is designed to be largely unaffected by random errors and delineations which give rise to outlying area measurements.
The method is based on the principle of 'vote taking'. It counts the number of times each pixel is included in a wound definition by all delineators, using binary mask images produced by filling the delineated boundary images. Thus a pixel which is never included by any delineator will have a score (vote) of zero, whereas a pixel which is always included by all delineators will have a maximum score which is set by the total number of area measurements and hence mask images. Pixels with a score of one are least-probable wound pixels, and pixels with a maximum score are most-probable wound pixels. Far outside of the wound the pixel scores will be zero. The centre of the wound should be a region of maximum score pixels. Near the boundary of the wound the scores of the pixels will lie within this range and the problem becomes one of setting a 'likelihood' threshold to discriminate between less likely and more likely wound pixels. A simple criterion used here is to select as wound pixels only those which have a score of at least 50%. This ensures a degree of robustness in the result since neither a rigid intersection of sets nor a loose union of sets will be robust in the presence of random errors, the union over-estimating the area and the intersection underestimating the area. Also, in the cases of single or infrequent events the union will be affected by outward placement of the boundary and the intersection by the inward placement of the boundary. The summation $R(x,y)$ of $N$ binary regions $P_n(x,y)$ is given by:

$$R(x,y) = \left\{ r(x,y) : N \bigg| r = \sum_{i,j \in \{1..N\}} (P_i(x,y) \cap P_j(x,y)) \cap (x,y) \in P_U(x,y) \right\}$$

where $P_U(x,y) = \bigcup_{n=1}^{n=N} P_n(x,y)$ i.e. the set of all pixels marked as wound pixels.

The median region $M(x,y)$ is the thresholded set of pixels, where $t$ is the threshold:

$$M(x,y) = \left\{ m(x,y) : r(x,y) \geq t = \frac{N+1}{2} \right\}$$
Figure 4.1 demonstrates the practical formation of the median region. The median region is formed when \( t = \frac{(N+1)/2} \) as in (4.1.7). The general intersection and general union of sets in Figure 4.1 are formed by setting \( t \) in (4.1.7) to \( N \) and 1 respectively. With this method one can include all 25 delineations for each wound to produce the most likely wound boundary, and one should expect its included area to be near to the median area of the 25 area measurements. The median boundary will be based on only five measurements when highlighting the differences between the boundaries delineated by the two delineators who produce the lowest and highest mean area measurements. The results in this case will be less robust an estimate, but will still allow the regions of ambiguity to be identified.

### 4.1.4 Model for Delineator Precision

The fractional precision estimate for a sample of wound area measurements is defined as

\[
PF = \frac{\sigma}{A}
\]

(4.1.8)

where \( \sigma \) is the area measurement precision and \( A \) is the ‘true’ area of the wound.

A general fractional precision estimate for a wound’s area may be defined as the arithmetic mean of each delineator’s precision values:
The fractional precision estimates of (4.1.8) and (4.1.9) are evidently quotients of two random variables. Linear statistical theory presents no exact formulae for calculating the expected mean value and standard error of such expressions. However, provided two variables are uncorrelated, it can be shown that:

\[
\begin{align*}
(4.1.10) & \quad \mathbb{E}\left[\frac{s_i}{\bar{a}_i}\right] = \frac{\mathbb{E}[s_i]}{\mathbb{E}[\bar{a}_i]} \\
(4.1.11) & \quad \mathbb{V}\left[\frac{s_i}{\bar{a}_i}\right] \approx \mathbb{V}[s_i]^2 \mathbb{V}\left[\frac{1}{\bar{a}_i}\right] + \frac{1}{\mathbb{E}[\bar{a}_i]^2} \mathbb{V}[s_i]
\end{align*}
\]

where \( \mathbb{E}\{\} \) and \( \mathbb{V}\{\} \) are the expectation and variance operators respectively.

The approximation in (4.1.10) relies upon \( \sigma^2/n \) where \( n \) is the sample size. Barford (1985) and Topping (1972) provide descriptions of combinations of standard error.

Assuming the area measurements produced by each delineator \( i \) are Gaussian with distribution \( a_{ij} \sim \mathcal{N}(A_i, \sigma_i^2) \) then the standard deviation estimate \( s_i \sim \chi_v \) where \( v=n-1 \) and the mean area estimate \( \bar{a}_i \sim \mathcal{N}(A_i, \sigma_i^2/n) \). It is now possible to define (4.1.10) in terms of measurable quantities:

\[
(4.1.12) \quad \mathbb{E}\left[\frac{s_i}{\bar{a}_i}\right] \approx \frac{\mathbb{E}[s_i]}{\mathbb{E}[\bar{a}_i]} = \frac{c_4 \sigma_i}{A_i}
\]

The standard square error of \( s_i/\bar{a}_i \) may now be expressed as:

\[
(4.1.13) \quad \mathbb{V}\left[\frac{s_i}{\bar{a}_i}\right] \approx \mathbb{V}[s_i]^2 \mathbb{V}\left[\frac{1}{\bar{a}_i}\right] + \left[\frac{1}{\bar{a}_i}\right]^2 \mathbb{V}[s_i] = \frac{c_2^2 \sigma_i^4}{nA_i^2} + \frac{\sigma_i^2 (1-c_4^2)}{A_i^2}
\]
where \[ \mathcal{V}\left\{ \frac{1}{\bar{a}_i} \right\} \approx \frac{\sigma_i^2}{A_i^4} \] and \[ \mathcal{V}\{s_i\} = \sigma_i^2(1-c_4^2) \]

Equations (4.1.12) and (4.1.13) together define the expected mean and expected standard error of estimate for the fractional precision definition of (4.1.8). The expectation of the estimate of fractional precision given by (4.1.12) is clearly biased with respect to (4.1.8) by the factor \( c_4 \). To produce an unbiased estimator of fractional precision it is necessary to define the estimator of fractional precision as:

\[
(4.1.14) \quad pf_i = \frac{s_i/c_4}{\bar{a}_i}
\]

Thus the expectation and variance of \( pf_i \) become, respectively:

\[
(4.1.15) \quad \mathbb{E}\{pf_i\} = \frac{\sigma_i}{A_i}
\]

\[
(4.1.16) \quad \mathcal{V}\{pf_i\} = \frac{\sigma_i^2(1-c_4^2)}{A_i^2c_4^2} + \frac{\sigma_i^4}{A_i^4n}
\]

The second term in (4.1.16) may be omitted provided \( A_i^4n \gg \sigma_i^4 \) yielding:

\[
(4.1.17) \quad \mathcal{V}\{pf_i\} \approx \frac{\sigma_i^2(1-c_4^2)}{A_i^2c_4^2}
\]

It is a straightforward matter to average the fractional precision estimates produced by each delineator for a given wound. Substituting (4.1.14) into (4.1.9) yields

\[
(4.1.18) \quad PF_{av} = \frac{1}{N} \sum_{i=1}^{i=N} \frac{s_i/c_4}{\bar{a}_i}
\]
The standard square error of $PF_{av}$ estimated using (4.1.17) is given by

\[(4.1.19) \quad \text{Var}\{PF_{av}\} \equiv \frac{1}{N^2} \sum_{i=1}^{i=N} s_i^2 (1 - c_4^2) \frac{1}{\bar{a}_i^2 c_4^2} \]

**Use of Pooled Variance Estimates**

The expected standard error of $PF_{av}$ can be improved and its calculation simplified when the delineator variances, $\sigma_i^2$, are equal for all delineators (for a given wound). The variance estimates, $s_i^2$, can thus be pooled increasing accuracy in estimating $\sigma^2$. The pooled variance equivalent of (4.1.18) becomes:

\[(4.1.20) \quad PF_{pooled} = \frac{\sqrt{MSE}}{Nc_4} \sum_{i=1}^{i=N} \frac{1}{\bar{a}_i} \text{ where } MSE = \frac{1}{N} \sum_{i=1}^{i=N} s_i^2 \]

The common number of degrees of freedom for $c_4$ and $MSE$, $v_{pooled} = Nv$

Using (4.1.17), the square of the standard error of $PF_{pooled}$ becomes:

\[(4.1.21) \quad \text{Var}\{PF_{pooled}\} \equiv \frac{MSE}{N^2} \left( \frac{1 - c_4^2}{c_4^2} \right) \sum_{i=1}^{i=N} \frac{1}{\bar{a}_i^2} \]

Note that (4.1.20) does not assume that the distributions from which the samples are drawn have the same mean value. This anticipates the existence of mutual biases among the set of experimental delineators.
4.1.5 Experimental Results

In total, the five delineators produced 250 area measurements upon which the following statistical tests are performed. The raw area measurements for each image were used to estimate the mean and variance of area measurements for each delineator upon each wound image. The mean and variance estimates for each wound and delineator are presented, respectively, in Tables B.1 and B.2, Appendix B.

4.1.5.1 Summary of Pre-ANOVA Check Results

There is no serious departure from the normality assumption in the cases of all 10 wounds. Figure 4.2 shows a typical plot of the residuals (taken from image 5), with one outlier marked. The probabilities are regressed twice upon the residuals, once with the outlier included and once with it excluded, showing the effect this outlier has upon the linearity of the plot. No departure from normality is apparent in this plot.

With most of the wounds, delineator variances do not differ significantly. In the case of wound 8, only one delineator produced an anomalous variance, which is a result of an altering opinion rather than variation caused by the limitations of manual dexterity. This analysis is covered in more depth in the next section.
A total of four outliers are detected in the measurements made by the delineators for the
ten wound images, and each case is considered to be due to a one-off change of
delineator judgment. The effect of these upon the mean and sum-of-square calculations
integral to the ANOVA tests is quite small, however.

4.1.5.2 Variance Analysis

The results of applying the Burr-Foster Q-test (4.1.2) for equality of several variances to
the area measurement data for each wound image are shown in Table 4.1, along with the
significance of the test in each case.

<table>
<thead>
<tr>
<th>Image</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q*</td>
<td>0.239</td>
<td>0.261</td>
<td>0.319</td>
<td>0.335</td>
<td>0.514</td>
<td>0.390</td>
<td>0.281</td>
<td>0.709</td>
<td>0.221</td>
<td>0.260</td>
</tr>
<tr>
<td>Signf.</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>&lt;1%</td>
<td>No</td>
<td>No</td>
<td>&lt;0.1%</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 4.1 Burr-Foster Q* Values for Equality of Variance Tests.

From tables: Q(α=0.010; N=5, v=4) = 0.443
Q(α=0.001; N=5, v=4) = 0.552

4.1.5.3 Bias Analysis

Table 4.2 shows the calculated ANOVA F-statistic for each wound, along with its
associated ‘P*’ value derived from the F-distribution with (4,20) degrees of freedom.
The null hypothesis that all five delineators measure the same area was rejected in the
case of each wound at least at the 5% significance level. Since this is the case for all 10
wounds the data are further examined to establish the differences between delineators. A
suitable test for this purpose is the Tukey-Kramer multiple comparison procedure (Neter
et al., 1996). Figure 4.3 shows a plot of these comparisons among the delineators with
95% confidence intervals for each wound. The maximum difference between the means
of the delineators – for each wound in turn – is calculated by:
\[ D_{\text{max}} = \max_i (\bar{a}_i) - \min_i (\bar{a}_i) \]

Table 4.2 contains this metric for each wound along with the associated Tukey-Kramer confidence intervals. The \( D_{\text{max}} \) metric is divided by the grand mean for each wound to provide an approximation to the fractional bias between the delineators who produce the most extreme area measurements.

<table>
<thead>
<tr>
<th>Image</th>
<th>A</th>
<th>( F_{\text{calc}} )</th>
<th>( P^* ) (%)</th>
<th>( D_{\text{max}} )</th>
<th>( D_{\text{max}}/A ) (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8458</td>
<td>5.03</td>
<td>0.6</td>
<td>580 ± 532</td>
<td>6.9 ± 6.3</td>
</tr>
<tr>
<td>2</td>
<td>6490</td>
<td>10.35</td>
<td>&lt;0.1</td>
<td>499 ± 245</td>
<td>7.7 ± 3.8</td>
</tr>
<tr>
<td>3</td>
<td>46194</td>
<td>4.14</td>
<td>1.3</td>
<td>1525 ± 1323</td>
<td>3.3 ± 2.9</td>
</tr>
<tr>
<td>4</td>
<td>35747</td>
<td>8.99</td>
<td>&lt;0.1</td>
<td>2692 ± 1573</td>
<td>7.5 ± 4.4</td>
</tr>
<tr>
<td>5</td>
<td>29152</td>
<td>8.94</td>
<td>&lt;0.1</td>
<td>5413 ± 2968</td>
<td>19.0 ± 10.2</td>
</tr>
<tr>
<td>6</td>
<td>24161</td>
<td>5.19</td>
<td>&lt;0.1</td>
<td>2246 ± 1715</td>
<td>9.3 ± 7.1</td>
</tr>
<tr>
<td>7</td>
<td>4072</td>
<td>9.51</td>
<td>&lt;0.1</td>
<td>633 ± 327</td>
<td>16.0 ± 8.0</td>
</tr>
<tr>
<td>8</td>
<td>38033</td>
<td>8.68</td>
<td>&lt;0.1</td>
<td>4478 ± 2617</td>
<td>12.0 ± 6.9</td>
</tr>
<tr>
<td>9</td>
<td>42577</td>
<td>23.88</td>
<td>&lt;0.1</td>
<td>7696 ± 3202</td>
<td>18.0 ± 7.5</td>
</tr>
<tr>
<td>10</td>
<td>43940</td>
<td>42.74</td>
<td>&lt;0.1</td>
<td>11252 ± 2936</td>
<td>26.0 ± 6.8</td>
</tr>
</tbody>
</table>

Table 4.2 Salient results of \( F \)-test for delineator agreement and subsequent bias estimation.

From statistical tables: \( F(1-\alpha = 0.95 ; 4, 20) = 2.87 \); \( F(1-\alpha = 0.99 ; 4, 20) = 4.43 \)

Further to the question of estimating the magnitude of the bias differences is the question of which particular regions of each image are the cause of bias can be answered by inspecting the average differences in boundary positioning achieved by the delineators who generated the extreme values of area. This is the Median Boundary method introduced in §4.1.3.3 and will show if the biases occur because certain parts of the image are excluded from the wound in a generally consistent manner or whether the regions are distributed in thin bands around the boundary of the wound.
Figure 4.3  Mean areas and Tukey HSD confidence intervals for delineated area bias analysis.
Figure 4.4 is a scatter-plot showing the correlation between the estimate of $D_{max}$ values of Table 4.2 and the difference of the areas produced by applying the median boundary method to those same samples. Clearly, in each case there is good agreement between the estimates of average wound area.

Figure 4.5 shows the gray images of the 10 wounds overlaid with the boundaries of the major regions of the differences detected by the median boundary analysis. Note that in most of the images the bias differences between the measurements of the delineators who produced the minimum and maximum mean area estimates is due to the formation of clustered regions of pixels. However, with some wounds, the pixels which cause the bias are scattered around the edge of the wound rather than forming clusters.

### 4.1.5.4 Test for Overall Delineator Bias

The previous hypothesis tests for bias shows that there is at least one delineator in the case of each wound whose measurements are biased. This being the case, it is possible to perform an overall significance test for bias which shows whether one or more delineators generally tend to produce area measurements with lower or higher values in relation to their peers. Since the area measurement variances among wounds are different, two-way ANOVA using the $F$-test is inappropriate. An alternative is to use a non-parametric test for differences, the Friedman test (Neave and Worthington, 1988).
Figure 4.5 Regions of significant differences identified from median boundaries
The Friedman 'S' statistic is approximately $\chi^2$ distributed with, in this case, 4 degrees of freedom.

Let $m_{wi} = \bar{a}_i$ where $\bar{a}_i$ is the mean area measurement produced by the $i$th delineator on wound $w$ where $w \in \{1..10\}$. The data model for the test with no interaction effects between the factors is defined as

$$m_{wi} = A_w + B_i + \epsilon_{wi}$$  \hspace{1cm} (4.1.23)

where $A_w$ is the area of wound $w$ and $B_i$ is the bias of delineator $i$.

The null hypothesis, $H_0$, is defined along with its alternative as:

$$H_0 : \text{All } B_i \text{ are equal}$$
$$H_1 : \text{Not all } B_i \text{ are equal}$$  \hspace{1cm} (4.1.24)

The differences between the true areas of each wound, $A_w$, are accounted for by blocking the data, i.e. each wound is defined as a block and the data, $m_{wi}$, are thus ranked within each block, removing the effect the different $A_w$. The rank sums for each delineator taken over the ten wounds are shown in Table 4.3 together with the value of the associated Friedman test statistic, $S$.

<table>
<thead>
<tr>
<th>Delineator, $i$</th>
<th>C</th>
<th>A</th>
<th>B</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank Sum</td>
<td>17</td>
<td>19</td>
<td>34</td>
<td>35</td>
<td>43</td>
</tr>
<tr>
<td>Friedman Statistic, $S$ :</td>
<td>18.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$P^*$ : Probability in tail area $\chi^2 &gt; 18.1$ :</td>
<td>0.0003</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.3 Rank sums and Friedman statistic for general delineator bias test.

$\chi^2$ Critical Value @ 1% significance level = 13.3.
As Table 4.3 shows, H₀ is rejected and therefore the conclusion is that delineators have a tendency to produce biased estimates on all wounds tested.

4.1.5.5 Precision Estimates

Table 4.1 shows that not all of the delineator variances for images 5 and 8 are equal. The pre-ANOVA checks identify ‘outliers’ for these images, but these outlying observations must not be discarded since they are not measurement errors. In this case, the averaged fractional precision formula given by (4.1.17) is to be preferred to the pooled-variance version of (4.1.19). The average fractional precision estimates for the area measurements of the 10 wound images are shown in Table 4.4. Fractional precision results for the area measurements are given using both the average-precision formula and the pooled-variance fractional precision estimates. As the table shows, the pooled variance formula produces somewhat diverse results for images 5 and 8 which have significantly differing variances among delineators.

<table>
<thead>
<tr>
<th>Image</th>
<th>$PF \pm \sigma(PF)$ (%)</th>
<th>$PF \pm \sigma(PF)$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><em>Pooled</em></td>
<td><em>Average</em></td>
</tr>
<tr>
<td>1</td>
<td>3.37 ± 0.24</td>
<td>3.46 ± 0.54</td>
</tr>
<tr>
<td>2</td>
<td>2.02 ± 0.14</td>
<td>2.05 ± 0.33</td>
</tr>
<tr>
<td>3</td>
<td>1.53 ± 0.11</td>
<td>1.49 ± 0.25</td>
</tr>
<tr>
<td>4</td>
<td>2.36 ± 0.17</td>
<td>2.29 ± 0.38</td>
</tr>
<tr>
<td>5*</td>
<td>(5.47 ± 0.39)</td>
<td>4.82 ± 0.87</td>
</tr>
<tr>
<td>6</td>
<td>3.80 ± 0.27</td>
<td>3.56 ± 0.60</td>
</tr>
<tr>
<td>7</td>
<td>4.31 ± 0.30</td>
<td>4.16 ± 0.67</td>
</tr>
<tr>
<td>8*</td>
<td>(3.69 ± 0.26)</td>
<td>2.99 ± 0.63</td>
</tr>
<tr>
<td>9</td>
<td>4.05 ± 0.29</td>
<td>4.19 ± 0.65</td>
</tr>
<tr>
<td>10</td>
<td>3.65 ± 0.26</td>
<td>3.74 ± 0.61</td>
</tr>
</tbody>
</table>

Table 4.4 Fractional precision estimates.  
* denotes differing variances.  
for $v=4$, $c_4=0.9400$.  
for $v=20$, $c_4=0.9876$.  

As Table 4.3 shows, H₀ is rejected and therefore the conclusion is that delineators have a tendency to produce biased estimates on all wounds tested.
4.1.6 Discussion

Bias Analysis

Table 4.2 shows that for each wound, the F-test for delineator bias (4.1.1) rejected the null hypothesis that all delineators measure the same area. Significant differences in area measurements therefore exist among delineators, so that for each wound image at least one pair of delineators disagreed in their average measurement of wound area. The Tukey confidence intervals shown in Figure 4.3 illustrate that the mutually biased area measurements from different delineators often fall into two or three overlapping groups. The area measurements for the three wound cases which produce the largest biases, images 5, 9 and 10, comprise two distinct non-overlapping groups in each case. This is also the case for the measurements of image 8 (6th largest bias). Except for image 9, the distinction is due to just one delineator in each case producing significantly smaller wound area measurements.

The regions of each wound which have been ambiguously interpreted, and thus are the cause of the biases, are shown in Figure 4.5. Bias arises in most cases when compacted regions of the wound, as opposed to elongated regions, are defined as wound by one delineator and as surrounding tissue by another. This is most clear in images 5 and 7 to 10 which have the largest biases.

The measurements for two of the wounds, images 5 and 8, both contain a single outlier which has increased the variances of the respective samples, thus causing the Burr-Foster 'Q' test to indicate significantly differing variances (Table 4.1). Inspection of the delineated area bitmap images shows that the outlier represents a change of opinion on the part of the delineator in each case. This will have inflated the mean-square-error (MSE) calculation for the bias analysis, causing wider confidence intervals. In the case of image 5, the delineator who produced the outlier was delineator A. The relevant line-plot in Figure 4.3 shows that delineator A is not one of the those who produced an extreme value of mean area. If the outlying observation is tentatively removed, the mean
area for delineator \( A \) increases but is still not an extreme mean area, thus the estimates of \( D_{\text{max}} \) in Table 4.2 are unaffected, except that, as mentioned, the confidence intervals are broader. In the case of image 8, delineator \( A \) again produces the outlier which inflates the sample variance and thus the MSE calculation. Delineator \( A \)'s mean area measurement is the lowest of the five delineators and the outlier is a lower-than average measurement for this delineator. Thus the \( D_{\text{max}} \) estimate is affected by the outlier.

The estimates of mean bias in Table 4.2 generally agree with the subjective three-group wound image classification described in §4.1.2. Biases range from 3.3% to 7.7% for the first group (images 1 to 3) and are generally lower than those in the other two groups and consistently lower than those in the most indistinct group (images 8 to 10) which has a bias range of 12.0% to 26.0%. This assumes that bias is independent of image size, or at least that it is fairly insensitive to it. The second experiment, presented in §4.2 below, will show that there is a bias component to manually-delineated area measurements, although it is small (approx. 2% maximum) in relation to the largest biases of the group containing images 8 to 10.

The Friedman test for bias across the whole wound-set (§4.1.5.4) indicates that the general differences of opinion which exist do so to some extent because of relative conservative and liberal attitudes on behalf of individual delineators, i.e. the magnitude of the overall differences between delineators is significant enough to show that delineators can be biased more often than not in one direction. Thus, some delineators generally measure higher or lower values of area in comparison with others. It is clear from the ranks in the above table that the five delineators fall into two distinct groups, with delineators \( A \) and \( C \) forming the lower group (under-estimators) and the other three delineators forming the upper group (over-estimators). It is possible to apply an analytical comparison procedure to identify significant differences, such as Dunn's paired comparisons (Neave and Worthington, 1988), although in this instance the differences are evident from the rank sums in Table 4.3. The distinction between the two groups arises because these two delineators frequently exclude parts of the image from
their wound definitions which are included by the others, although they were not always responsible for generating the lowest mean area measurements for each wound.

**Precision Analysis**

Table 4.4 shows that the fractional precision of area measurements varies, depending upon the particular wound, from 1.53±0.11% for image 3 (pooled variance) up to 4.82±0.87% for image 5, although it is not possible to state how much of this variation is due to the size of the wound (in pixels). The inflationary effect of unequal variances upon the pooled-variance version of the precision estimate is clearly seen in the disparity between the entries for the two versions of estimated precision for images 5 and 8. Otherwise, the agreement between the pooled estimate version (equation 4.1.18) and the averaged version (equation 4.1.20) is evident. Note also that the associated standard error estimates when using pooled-variances are less than half the respective error estimates produced by the averaged version. In contrast to the case for bias there is no clear distinction between the precision values for the first three wounds (considered visually unambiguous) and the last three wounds which are considered visually ambiguous. Thus it appears that the effect of manual dexterity and the mechanical properties of the delineation equipment (in this case a mouse) have swamped any variation due to vagueness of the wound boundary.

4.1.7 Conclusions

- Delineators produce wound area measurements that are mutually biased and thus do not always agree on the average area of a wound. The relative size of such bias between delineators is dependent upon the particular wound. The estimated largest mutual bias existing between any two delineators ranges from 3.3±2.9 % to 26.0±6.8 %.
Delineators have a general tendency to produce area measurements which on average are biased in one particular sense, i.e. comparison between different delineators’ wound area measurements over a collection of wound images shows that some delineators tend to under-estimate the area and others to over-estimate the area.

Subjective assessment of a wound image can be indicative of the amount of bias expected between delineators, since bias is dependent upon the level of ambiguity that exists relating to the boundary of the wound. This greater levels of bias quoted above are probably large enough to give cause for concern when comparing measurements made on consecutive hospital/home visits unless the measurements are made by the same staff member. A further study is necessary if one wishes to measure the rate of change of wound for various wounds. For instance, it may be found that the area of a healing wound diminishes with a certain characteristic function.

Equal levels of area measurement precision generally exist among delineators (8 cases out of 10), although delineators tend to exhibit relatively poor precision when they alter their opinion upon which areas of an image belong to a wound (images 5 and 8). In contrast, the fractional precision does vary from wound to wound with the average fractional precision estimate for wounds ranging from 1.53±0.11 % to 4.82±0.87%. It is not clear, however, whether this difference of precision among wounds is a function of the ambiguity and relative vagueness or clarity of the wound boundary in an image. The images were of greatly varied sizes and so it may be argued that precision is affected by image size.
4.2 The Effect of Image Display Size upon Delineation Accuracy

4.2.1 Objectives

This experiment aims to determine the following:

- Whether or not the variation inherent in the manual delineation process is independent of the size of the image displayed and is therefore constant over the range of sizes used for the experiment.

- Whether or not a bias exists between measurements taken at different levels of scale.

It is desired to express the results concerning both of these objectives in terms of percentages of wound area in order to aid appreciation of the sizes of errors concerned. In accordance with the stated objectives, the data collected throughout this experiment will be used in the fitting of two regression models, one for bias and the other for variation, both of them expressed as functions of display scale.

4.2.2 Development of Regression Model for Bias

The proposed model for bias is developed here with reference to a practical delineation error model which is shown in Figure 4.6(a). This model seeks to express the bias error as a consequence of the discretisation of the image plane into finite size pixels. In Figure 4.6(a), delineation of the wound boundary is shown as a finite-thickness line centred at the wound edge. Here, the drawing implement is modelled as a mouse pointer drawing a line of 1 pixel thickness. It is the thickness of this line (and its length, i.e. the perimeter-length of the shape) in relation to the actual wound size that governs the bias error, the error being the area enclosed between the outer edge of the thickened delineating line and the centreline. Thus it is possible to intuitively write an equation expressing this
notion and attempt to validate it. A first approximation to a working model for bias is given by (4.2.1).

\[
A(\lambda) = \lambda^2 A_0 + \lambda B_0 = \lambda^2 A_0 + \lambda P_0 \delta
\]

where \( B_0 = P_0 \delta \) is the bias coefficient
\( P_0 \) is the perimeter-length of the boundary
\( \delta \) is the delineation translation error constant
\( \lambda \) is the linear magnification factor.

This equation has two components: The first is a proportional term in \( \lambda^2 \) having slope \( A_0 \) which equates to the true area of the wound given the area magnification factor, \( \lambda^2 \). The second is a component proportional to \( \lambda \) which equates to the absolute size of the bias error, and is a product of perimeter length \( P_0 \) and a constant offset, \( \delta \). In relation to the scaled true-area component, the bias component diminishes with increasing scale, and hence in general the curve will tend to appear quite linear, since one expects \( A_0 \gg B_0 \).

To aid appreciation of the nature of the bias component it is possible to express – at any
scale – bias as a fraction of the scaled area, by dividing both sides of (4.2.1) either by the area component, $\lambda^2 A_0$, of (4.2.1) or simply by $\lambda^2$. Equation (4.2.2) is the result of dividing through by $\lambda^2$ which yields the expected prototype plot shown in Figure 4.6(b) – a constant term plus inverse-square root term decaying from a highly biased position at low levels of the predictor variable, $\lambda^2$. As the predictor variable increases towards infinity, the response curve asymptotically approaches the horizontal line drawn at the true value of the variable being measured. This plot is included here to aid appreciation of the expected nature of the bias function and of the coefficients $A_0$ and $B_0$. The experimental results for the bias analysis will be shown in a form similar to Figure 4.6(b).

\[
N(\lambda) = \frac{A(\lambda)}{\lambda^2} = A_0 \left( 1 + \frac{B_0}{A_0 \lambda} \right)
\]

Although (4.2.1) is not linear in $\lambda$, its normalised form (4.2.2) may be considered linear in a variable $1/\lambda$. This form of presentation would show the fractional bias as a function of scale. Applying the linearised model to the area measurement data yields residuals which are not particularly normal, have increasing variance and skewed spacing of the predictor levels. In contrast, it has been found that working with a basic datum related to the square root of the area produces residuals which better satisfy the assumptions of normality and constant variance. As part of the validation of (4.2.1), consider the bias in area measurements that would arise from delineating scaled versions of a circle having radius $r_0$ with a constant offset $\delta$. This is visualised in Figure 4.6(a) and the effect is modelled by (4.2.3):

\[
r(\lambda) = r_0 \lambda + \delta
\]

Using the standard formula $A_{\text{circle}} = \pi r^2$ and substituting (4.2.3) for $r$ yields:

\[
A_{\text{circle}}(\lambda) = \pi (r_0 \lambda + \delta)^2 = \pi r_0^2 \lambda^2 + 2 \pi r_0 \delta \lambda + \delta^2
\]
Identifying the individual terms, the equation may be expressed as:

\begin{equation}
A_{\text{circle}}(\lambda) = \lambda^2 A_0 + \lambda C_0 \delta + \delta^2
\end{equation}

where area $A_0 = \pi r_0^2$ and circumference $C_0 = 2\pi = P_0$.

Note the similarity between the expressions for delineated area in (4.2.1) and (4.2.5). Equation (4.2.5) has an additional term, $\delta$, which is negligible when $\delta << r$.

Equation (4.2.5) is an exact relationship for the delineated area when the object is a circle and therefore the intuitive model of (4.2.1) is inexact for a circle because of the omission of the constant $\delta$, which is small, however, when $\delta << r$, which is generally the case. The final step in determining a bias model expressed in radial terms is to generalise (4.2.3) to allow it to apply to non-circular objects. Equation (4.2.6) defines the 'average radius', $\rho$, of a general shape having area $A$:

\begin{equation}
\rho = \sqrt{A/\pi}
\end{equation}

The 'average radius' $\rho$ is thus, equivalently, the radius of a circle having area equal to $A$.

For generalised shapes, (4.2.7) defines the scale-invariant shape factor $F$, for a simply-connected object (Wahl, 1987), where $P$ is the perimeter length of the object and $A$ its area. When the shape is a circle, $F$ has a value of unity, for non-circular shapes $F > 1$.

\begin{equation}
F = \frac{P^2}{4\pi A} \quad \text{thus} \quad A = \frac{P^2}{4\pi F}
\end{equation}

If $C$ is the circumference of a circle with area $A$, and $P$ is the perimeter-length of a general shape also with area $A$ and $F$ is its scale-independent shape-factor, then equating the expressions for area yields:
(4.2.8) \[ A = \frac{P_{\text{shape}}^2}{4\pi F_{\text{shape}}} - \frac{C^2}{4\pi} \]

Thus the perimeter of the shape, which partly governs the area bias, is given by:

(4.2.9) \[ P_{\text{shape}} = C \sqrt{F_{\text{shape}}} = 2\pi p \sqrt{F_{\text{shape}}} \]

Substituting for \( P_{\text{shape}} \) from (4.2.9) into (4.2.1) yields:

(4.2.10) \[ A(\lambda) = \lambda^2 A_0 + \lambda \sqrt{F_{\text{shape}}} C_0 \delta \]

Where the factor \( F_{\text{shape}} \) provides the enlargement of the bias term above that expected from a circle of equal area by virtue of the shape’s elongated perimeter length in relation to its area. Applying this notion to the radial model of (4.2.3) and thus generalising it to apply to non-circular shapes gives a linear equation in \( \lambda \):

(4.2.11) \[ \rho(\lambda) = \rho_0 \lambda + \sqrt{F} \delta = \beta_1 \lambda + \beta_0 \]

Thus the simple linear regression model for noisy observations of \( \rho(\lambda) \) becomes:

(4.2.12) \[ \hat{\rho}_i = \hat{\beta}_0 + \hat{\beta}_1 \lambda_i + \varepsilon_i \] where \( \varepsilon_i \sim \text{NID}(0, \sigma^2) \)

For each set of results, this equation may be used to obtain estimates \( b_0 \) and \( b_1 \) respectively of the regression parameters \( \beta_0 \) and \( \beta_1 \). From (4.2.11) the following identities may be established:

(4.2.13) \[ b_0 = \hat{L} \delta \] from \( \beta_0 = \sqrt{F} \delta \) where \( L = \sqrt{F} \)

(4.2.14) \[ b_1 = \hat{\rho}_0 \] from \( \beta_1 = \rho_0 \)
From these two identities one can obtain the following (biased*) estimates:

\[(4.2.15) \quad \hat{A}_0 = \pi \hat{b}_1^2 \quad \text{the true (unbiased) area}\]

\[(4.2.16) \quad \hat{\delta} = \frac{b_0}{\hat{L}} \quad \text{the delineation translation error constant.}\]

* The use of the term 'biased' here refers to the fact that a square-transformation of an unbiased estimate introduces a small bias into the estimate of the transformed variable by virtue of its 'skewing' effect upon the distribution of the transformed variable. However, it is the case that when the value of the estimate is much greater than its standard error, the magnitude of the skew becomes negligible, as in this case. Using the expression for average radius, \(\rho(\lambda)\) from (4.2.11), the equivalent expression for delineated area becomes:

\[(4.2.17) \quad A(\lambda) = \pi (\rho_0 \lambda + L \delta)^2 \equiv \hat{\lambda}^2 A_0 + \hat{\lambda} B_0 \]

where \(A_0 = \pi \rho_0^2\) and \(B_0 = 2\pi \rho_0 L \delta = \rho_0 \delta\)

\(\pi \delta^2\) is the error introduced by the approximation.

Hence, the fractional area bias error may be expressed as:

\[(4.2.18) \quad BF(\lambda) = \frac{1}{\hat{\lambda}} \frac{B_0}{A_0} = \frac{2}{\hat{\lambda}} \frac{L \delta}{\rho_0} \]

In terms of the estimated regression parameters, (4.2.18) becomes:

\[(4.2.19) \quad BF(\lambda) = \frac{2}{\hat{\lambda}} \frac{b_0}{\hat{b}_1} \]

The standard errors of the regression parameters, \(s\{b_0\}\) and \(s\{b_1\}\), are obtained from the regression analysis and as such are automatically produced by standard statistical
packages – see Neter et al. (1996) for the requisite formulae. In the present case the parameters of interest are not $b_0$ and $b_1$ directly, but rather functions of them, viz.: equations (4.2.15), (4.2.16) and (4.2.19). When functions of random variables are non-linear in form, the standard error of the function result cannot be exactly expressed in terms of the standard errors of its arguments. Generally, the standard error, $S$, of any combination of independent estimates ($m_1, m_2, ..., m_n$) with respective standard errors ($S_1, S_2, ..., S_N$) may be first-order approximated by:

\begin{equation}
S^2 \equiv \sum_{i=1}^{N} \left( S_i \frac{\partial f}{\partial m_i} \right)^2
\end{equation}

In each specific applied case of the general function, $f$, from (4.2.20), the appropriateness of the approximation should be considered. Applying (4.2.20) to the estimation of unbiased area $A_0$ (4.2.15) the standard error is approximated by:

\begin{equation}
s(A_0) \equiv 2n b_1 s(b_1) \quad \text{provided} \quad s(b_1) << b_1.
\end{equation}

Although $b_0$ and $b_1$ are not completely independent, the dependence should be negligible when the total sample size of the regression data is large (>50). Thus the standard error of the estimate of $BF$ from (4.2.19) should yield:

\begin{equation}
s(BF) \equiv \frac{2}{b_1} \sqrt{s^2(b_0) + \left( \frac{b_0}{b_1} \right)^2 s^2(b_1)} \quad \text{provided} \quad s(b_1) << b_1.
\end{equation}

Finally, (4.2.19) may be expressed as a function of the magnified unbiased area, so that the direct relationship between fractional bias, $BF$, and $A_0$ may be realised:

\begin{equation}
BF(\lambda; A_0) = \frac{2L \delta}{\lambda} \frac{\pi}{A_0}
\end{equation}
4.2.3 Development of Regression Model for Variance

The proposed model for variation of area measurements as a function of the area magnification factor, $\alpha$, is

$$\sigma^2 = k\alpha^n \quad \text{where} \quad \alpha = \lambda^2$$

$k$ is a coefficient taking into account the sum variation due to the particular wound image, delineator and equipment used.

Equation (4.2.24) may be linearised by taking a logarithmic transform of both sides, which allows the use of the simple linear regression model. Applying this transformation to (4.2.24) yields:

$$\log(\sigma) = \log(k) + n\log(\alpha) = \beta_0 + \beta_1 \log(\alpha)$$

where $\beta_0 = \log(k)$ and $\beta_1 = n$

It should be noted that under the logarithmic transformation of (4.2.25) the expected response curves for variance and standard deviation are linearly related, so that:

$$\log(\sigma) = \frac{1}{2} \log(\sigma^2) = \frac{1}{2} \log(k) + \frac{n}{2} \log(\alpha)$$

The fractional precision is given by the standard deviation of area measurements expressed as a fraction of the expected scaled area at scale $\alpha$. Thus using the hypothetical relationship between $\sigma^2$ and $\alpha$ from (4.2.24), one obtains:

$$PF = \frac{\sigma}{\alpha A_0} = \frac{\alpha^{n-1} \sqrt{k}}{A_0}$$
If magnification of an image has no effect upon the fractional precision, \( PF \), then in reality (4.2.27) must be a constant expression, independent of \( \alpha \). This requires that the parameter \( n=2 \), signifying that the standard deviation of area measurements is proportional to the area. In terms of the regression analysis this translates to a null hypothesis that \( \beta_1=2 \). This hypothesis is considered supported when the 95% confidence interval for \( \beta_1 \) overlaps 2. If the fractional precision improves with increasing magnification of the image, then one expects \( \beta_1<2 \).

Taking logarithms of both sides of (4.2.27) yields:

\[
\log(PF) = \frac{1}{2}\log(k) - \log(A_0) + (n/2 - 1)\log(s) = \beta_0^* + \beta_1^* \log(\alpha)
\]

where \( \beta_0^* = \frac{1}{2}\log(k) - \log(A_0) \) and \( \beta_1^* = \frac{n}{2} - 1 \)

Comparing (4.2.26) with (4.2.28) it is clear that they are linearly related, so that:

\[
\beta_0^* = \frac{1}{2}\beta_0 - \log(A_0)
\]

\[
\beta_1^* = \frac{1}{2}\beta_1 - 1
\]

These are the model parameters expressing the fractional precision of area measurements as a function of the area-magnification parameter \( \alpha \). Since \( A_0 \) is the only unknown variable in this transformation, the effect of the uncertainty introduced into the fitted precision model by using an estimate of \( A_0 \) is to add a small amount of variability into the intercept parameter (i.e. the logarithm of \( k \)). The slope parameter (exponent) is unaffected by such additional error.

When precision is expressed as a direct function of the scaled area of a wound \( (\alpha A_0) \), a meaningful comparison can be made between the precision curves estimated for different wounds produced by different delineators. This is possible because the
common factor for comparison of delineation performance is the actual display size of the wound.

4.2.4 Experimental Set-up and Procedure

Two volunteer delineators were asked to use a standard computer mouse to make repeated delineation of three different wounds, with each wound displayed at seven different sizes. The delineation software produced for the first experiment (§4.1) was used to perform the delineation task. The set of linear scale ratios was chosen to be $\Lambda = (1.00, 1.33, 1.66, 2.00, 2.33, 2.66, 3.00)$ and these values were used to scale both the horizontal and vertical dimensions of each of the three wound images. Thus for each wound set the ratio of the largest area to the smallest area is 9:1, since the area scale is the square of the linear scale. The largest image (area scale 9) in each set was scaled from its original digitised image so that its dimensions fitted a standard 800×600 pixel display. The smaller scaled versions were then produced with appropriate sizing relative to the largest version.

Three different wound images were selected from a library of wound images. The first image was image 1 from the first delineation experiment and was selected because it has a relatively well-defined boundary, is fairly homogenous and does not lead to large differences of opinion concerning the wound’s area (cf. §4.1.4.3). The second image, a ‘phantom’ wound modelled on the first image, was created using a mask defining the wound (cf. §4.2.4.1). The third wound image, image 6 from the first experiment, was selected because the wound boundary was subjectively less well defined than the first image. The reason for including such an image is to identify whether image quality/clarity affects bias and precision.

Both delineators were asked to delineate all seven versions of the three images once per day for ten days and record the result as a bitmapped image of edge pixels, as in the first
experiment. The area of each delineation was thus calculated by counting pixels both inside and on the pixelated boundary.

4.2.4.1 Creation of Phantom Wound Image

To create the phantom image a wound template was first obtained from the 25 delineations produced for image 1 from the first delineation experiment. The mask of the median boundary of the associated 25 wound masks was produced and contained 8779 pixels. Secondly, samples of the colour pixels from image 1 were taken both from inside and outside of the wound in order to more realistically colour the phantom image. Finally, the coloured phantom image (size: 256² pixels) was blurred by application of an isometric Gaussian smoothing filter with σ=1.5 pixels and the smoothed image was used to create the 7 scaled versions required for the experiment by bilinear interpolation. The mask image was also scaled to coincide with the smallest phantom image (α=1) and contained 5358 wound pixels. This is the true value of $A_0$ for this image.

The notion of using the phantom wound image is that it should enable the measurement of delineator performance on a wound shape without the influence of ambiguous boundary segments.

4.2.5 Bias Model Results

The average radius data $\rho_i$ was calculated from the measured area data with (4.2.6). Using the regression model of (4.2.12), the average radius data for each set of measurements was regressed upon the linear scale variable, $\lambda$, to provide parameter estimates $b_0$ and $b_1$. From these regression parameters, the values of $A_0$ and $BF_0$ were calculated using (4.2.15) and (4.2.19) respectively. The mean value of $L$ for each delineator sample was calculated from the delineated images and used to estimate $\delta$. 
Table 4.5 contains the estimates of $A_0$ and $BF_0$ together with the estimates of $L$ and $\delta$. The area-related parameters $A_0$ and $BF_0$ are used to plot the fractional bias against unbiased magnified area as shown in Figure 4.7(a)-(d) for all sets of results. The relationship between fractional bias and unbiased area used to plot these graphs is given by equation (4.2.19).

![Table 4.5](image)

Table 4.5 Parameter estimates and associated standard errors for bias analysis

Figure 4.8 plots 95% confidence intervals of the estimated area constant $A_0$ for images 1 and 2. Figure 4.9 plots 95% confidence intervals of fractional bias constant $BF_0$ for all three images. The area estimates from both delineators for images 1 and 2, which are essentially of the same wound, are combined produce the following higher-confidence estimates of $A_0$:

Area of image 1 : $(5230 + 5198)/2 = 5214 \pm 18$ pixels
Area of image 2 : $(5346 + 5351)/2 = 5349 \pm 10$ pixels
Difference \hspace{1cm} = 135 \pm 21 pixels
Figure 4.7 Transformed fractional bias plots showing bias as a function of unbiased magnified area for (a) delineator 1 and images 1 and 2, (b) delineator 2 and images 1 and 2, (c) delineator 1 and image 3, and (d) delineator 2 and image 3.

Figure 4.8 Confidence intervals for estimates of unbiased wound area, $A_o$, for images 1 and 2.

Figure 4.9 Confidence intervals and estimates of fractional bias constants, $BF_0$, for all three images.
4.2.6 Variance Model Results

For each wound/delineator the variance estimate of the area measurements at each level of scale was calculated. Using (4.2.23) the logarithm of the estimated area variance at each level of area scale was regressed against the logarithm of the area-scaling factor. The estimated regression parameters \( b_0 \) and \( b_1 \) are shown in Table 4.6 with estimated standard errors. The fitted linear regression curves obtained with these estimates are plotted in Figure 4.10(a,c,e). Each graph shows a comparison between the results obtained from the data produced by the two delineators. The linear equations were transformed by (4.2.24) to produce fractional variance results. Figure 4.10 (b,d,f) shows precision plotted as a function of wound area for each of the three images. Again, each graph compares the curves produced from the two delineators' results. The \( b_1 \) linear regression parameter determines the specific form of the variance/scale relationship. The estimates of \( b_1 \) obtained from the results of both delineators' measurements for each image are plotted in Figure 4.11 along with the appropriate 95% confidence intervals.

<table>
<thead>
<tr>
<th>Image</th>
<th>Delineator</th>
<th>( b_0 = \log k )</th>
<th>( b_1 = n )</th>
<th>( y_k )</th>
<th>( y_k/A_0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3.86, 0.11</td>
<td>1.79, 0.18</td>
<td>85</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.36, 0.17</td>
<td>0.87, 0.27</td>
<td>151</td>
<td>0.029</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>3.42, 0.08</td>
<td>1.63, 0.12</td>
<td>51</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.67, 0.13</td>
<td>1.47, 0.20</td>
<td>68</td>
<td>0.013</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>4.29, 0.17</td>
<td>1.46, 0.27</td>
<td>140</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4.27, 0.18</td>
<td>1.96, 0.29</td>
<td>136</td>
<td>0.025</td>
</tr>
</tbody>
</table>

Table 4.6 Parameter Estimates and Transformed Coefficients for Variance Regression Analysis
Figure 4.10 (a, c, e) Plotted regression of area variance on area magnification factor $\alpha$ for wound images (1, 2, 3) respectively (equation 4.2.25). Each graph compares the regression lines produced from the results of both delineators, (b, d, f) plot the corresponding transformed functions (equation 4.2.27) for images (1, 2, 3).
To make a comparison between the fractional precision levels for all three images, the estimated area constant, $A_0$, is used to determine two levels of the predictor variable, $a$, required in each bias function plotted in Figure 4.7 that sets the unbiased area at (a) 10,000 pixels and (b) 40,000 pixels. This involves using the regression function to predict the value of the response variable, $\log(s^2)$, at two new $a$ levels and transform them into approximate predictions of fractional precision. The results of the predictions for fractional precision of area measurements at the two specified sizes of wound area are shown in Table 4.7.

![Figure 4.11](image)

**Figure 4.11** Confidence intervals for area variance power constant for each image and for both delineators.

<table>
<thead>
<tr>
<th>Image</th>
<th>Delineator</th>
<th>$A_0$</th>
<th>$PF$ (%)</th>
<th>$PF$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>$A(a)=10000$ pixels</td>
<td>$A(a)=40000$ pixels</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>5214</td>
<td>1.5%</td>
<td>1.3%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5214</td>
<td>2.0%</td>
<td>0.9%</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5349</td>
<td>0.9%</td>
<td>0.7%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5349</td>
<td>1.1%</td>
<td>0.8%</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>6179</td>
<td>2.0%</td>
<td>1.4%</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>5548</td>
<td>2.4%</td>
<td>2.4%</td>
</tr>
</tbody>
</table>

**Table 4.7** Fractional precision estimates at specific areas
4.2.7 Discussion

A. Bias Analysis

Area Constant, $A_0$

The area constant, $A_0$, for the first two images, 1 and its phantom, 2, are all estimated with ±1% accuracy. Also, the estimates of $A_0$ for each image are comparable between the two delineators so that the average of the two areas for each wound provide a high-confidence estimate of $A_0$, respectively 5214±18 pixels and 5349±10 pixels which is subsequently required for the variance analysis. The difference between these values of 136±21 pixels is not expected and efforts to trace the source of this disparity have failed. The mask area (5358 pixels) was subsequently found to be larger than the scaled version of the median boundary mask by a similar amount, so the disparity may be explained as a lapse of procedure.

The estimates of area constant $A_0$ for image 2 (phantom) from both delineators were consistent with the scaled true value of the mask used to generate the phantom. This is a practical validation of the model for bias at least when using a phantom image.

Estimation of translation constant, $\delta$

The column for $\delta$ in Table 4.5 records the fact that four of the six estimates of the translation constant produced results about twice as large as expected, given the prior discussion in §4.2.2. The two estimates which support the default value of $\delta = \frac{\sqrt{2}}{2}$ are obtained from image 3 and have the widest confidence intervals, suggesting that since they are in agreement to some extent with the first four estimates of $\delta$ that there is no contradiction in concluding that $\delta > \frac{\sqrt{2}}{2}$. Since this value is greater than originally supposed, it does not invalidate the notion of the $\frac{\sqrt{2}}{2}$ pixel width translation being a factor in producing bias, but it is necessary to accept that the explanation is incomplete. Thus,
it is proposed that an extra translation component exists that increases the bias. One hypothesis for producing such an increased shift is to consider the influence of the display monitor’s γ-factor upon the apparent positioning of the edge of the wound. This would certainly seem to shift the supposed centre of an edge whose contrast extends over several pixel-widths. However under the model of (4.2.1) the constant serves to affect only the bias (μ) term. In other words, a translation effect that is proportional to the width of the edge will serve to magnify the area – affecting the λ² term in (4.2.1) – and will not affect the bias term. The enlargement of the translation constant could be explained if human visual perception somehow normalises the apparent width of an edge and subsequently assesses the edge of the wound as being offset by the displacement due to the monitor’s γ-factor. To be sure, it would be necessary to conduct a further study attempting to measure this effect uniquely. Regardless of this however, the magnitude of the bias effect upon area measurements has been successfully measured.

With reference to the bias curves shown in Figure 4.7, it can be seen that both delineators produce comparable results for bias when delineating image 1, with the bias contributing about 6% (i.e. reading 0.06 from the graph) to the area measurement at the lowest scale tested and 2% of the area measured at the highest scale. Comparing the bias results for images 1 and 2 clearly shows two things:

1) Delineator 1 recorded a reduction in bias when delineating the phantom image (2), with the level of bias introduced being roughly half that of the bias from image 1, while delineator 2 recorded no such improvement.

2) There is great similarity between delineator 2’s curves for images 1 and 2 and delineator 1’s curve for image 1, suggesting that the bias level is consistent between the two delineators, except for delineator 1’s attempt on image 1 which is much more consistent with delineator 1’s attempt on image 3.
B. Precision Analysis

The regression curves for variance in Figure 4.10(a) and (b) for image 1 show that delineator 1 is more precise at lower scales than delineator 2 but becomes less precise at higher scales. This comparison between the two delineators is not consistent across the three images and the disparity in performance between the delineators on images 2 and 3 is much less than that on image 1. The entries for the exponent parameter $n$ (delineator 2) in Table 4.2 (also displayed in Figure 4.11) show that the value of 0.87 is the lowest recorded. However, the size of the confidence intervals for all six measurements of $n$ imply that there may well be no difference of any practical size over both delineators and the three images, with a mean estimated value of $1.53 \pm 0.09$.

The range of the area magnification factor $A^2$ translates to an approximate range of 5000 pixels to 50000 pixels for the areas of the wounds measured. This translates again, approximately, to an actual display size of $5cm^2$ to $50cm^2$ (calculated from a 15” monitor, 800×600 pixel mode).

One factor that could affect variability is the time taken to delineate a wound. Thus it is possible to conjecture, for instance, that if one delineator takes roughly the same amount of time to delineate a wound, regardless of its size, then one could expect a steeper precision curve in comparison with a delineator who applied equal diligence to the delineation process - hence requiring more time to delineate the longer boundaries of larger wounds - regardless of scale.

The results for image 3 are less precise than the results for images 1 and 2 (the better-defined images). The precision curve for delineator 2 (Figure 4.10 (f)) shows no predictable change in precision over the experimental range, with the value of $n=2$. This delineator delineated a different boundary to delineator 1, choosing a rather difficult to delineate boundary section, which could account for unpredictable results in terms of variances.
4.2.8 Conclusions

- The experiment has clearly demonstrated that measurements of wound area made by manual delineation with a mouse on a pixelated display are subject to small bias errors. These errors are dependent on the actual size of the displayed image, upon the delineator and upon the wound image itself. The model for explanation of the production of bias (4.2.3) provided a larger than expected value for the delineation translation error (4.2.16) in each case, a result which is difficult to explain, but suggests that there is more than the one proposed factor causing bias. Nonetheless, the manifestation of delineator bias has been made known and its form and magnitude estimated for several cases.

- Since bias is deemed to exist and its characteristic function and (partial) causal effect is known, a correction for bias may be necessary when using manual area measurements for assessing the progress of wound healing. This may prove less than adequate in practice if it were subsequently to be found that bias varies between images and even delineators. Active contour models do not suffer from this form of bias because the contour is a line (of zero thickness) and is not constrained to a discrete grid (in the case of the finite element models).

- The experiment has shown that in general the fractional precision of wound area measurements made by manual delineation improves with increasing magnification of the displayed image. This result signals that when small wounds are imaged, or more generally, when the wound covers a relatively small number of pixels, that the variation in manual delineated area measurements will be adversely affected. Thus wound images should be properly scaled for display in order obtain higher precision.

- No distinct overall differences in precision are observable between the delineators and the general precision level varies from 1.5% to 3% at low scales (approx. 5 cm²) down to 0.8% to 1.5% at high scales (approx. 50 cm²), depending on the wound. In the case of delineator 2’s attempt at image 3 the variation introduced by repeatedly delineating a rather vague boundary portion tended to swamp the general variation due to the effect of manual dexterity.
PLATE 1  Wound images used for the first manual delineation performance study (§4.1), the active contour model parameter setting experiments and the final performance analysis.

Original in Colour
PLATE 2 Additional set of wound images used in the final performance study of the active contour model algorithms.

*Original in Colour*
5 PARAMETER SETTING EXPERIMENTS

5.1 Introduction and Preliminary Notes

The two finite element active contour models described in Chapter 3 – the Ground-Attract (GA) and Tangent-Normal (TN) algorithms – each have two internal parameters which govern the weights of the elastic and stiffness energy terms of the active contour model. The levels at which these parameters are set will affect each algorithm’s performance. Due to the differences in the formulation of the internal energy terms for the two algorithms, the behaviour of the contours will not be the same but will share some similar characteristics. The energy component weights of the two minimax algorithms (MX and MG) are implicitly set and so require no such concerns. However, in addition to the setting of internal parameters, the image itself is subject to regularisation (smoothing or low-pass filtering) to encourage uniqueness of the boundary solution under differing starting contours.

Regularisation of the image (external energy function) is a necessary component of all four algorithms previously described and it is necessary to determine, for each algorithm, the optimal filter scale parameters for the scale descent algorithm described in §3.1. The purpose of the experiments described in this chapter is to determine the effects of varying these parameters upon the area enclosed by the contour over a suitably wide range, given that the objective of the application of these algorithms is ultimately to reliably and accurately measure the wound area.

This chapter considers the effect of the following variables that have the ability to affect the final solution of the active contour models:

- Contour regularisation parameters.
- Image regularisation parameters.
- Initial position.
All three variables have the potential to influence the bias and precision of the measured area of a wound. It should be considered that each of the variables is either essentially a random variable (initial position – see below) or an unknown parameter which must be set at some appropriate level in order to produce desirable results. Certain parameters are left unaltered through the course of these experiments, viz.: number of contour elements (set at 32) and the step size reduction scheme (unit time or space steps, reducing in single octave steps until converged at 1/32).

5.1.1 A Note on Varying the Initial Contour Position

A spatial discretisation of the generalised parametric contour \( v(s) \) into a piecewise continuous form allows one to specify changes to the contour in a finite number of degrees of freedom. Clearly, in each image there is a region in the vicinity of the wound boundary within which a suitably regularised contour should converge to an approximation of the boundary. This corresponds to an 'energy well' surrounding a local minimum in the active contour model's energy functional. Ambiguous contours will affect the consistency of such a solution. However, there will be possible initialisations of the contour where the contour covers part of the image that contains edges other than those belonging to the wound which will lead to a boundary solution which does not uniquely apply to one object in the image. Defining this domain for any image would be a difficult task. The domain would be unique to each image and would be affected by the regularisation parameters. Rather than attempting to generate arbitrary initial contours that may produce hybrid solutions, the approach taken here is to use a set of manual delineations to initialise the active contour model. These delineations should not place the contour over a boundary segment other than those which belong to the wound, although ambiguity of the exact wound boundary will clearly lead to initialisations in different energy wells thus implying differing results.
5.1.2 Calculation of Contour Area

In the following experiments the basic datum is the area enclosed by a piece-wise parametric contour. All algorithms use the cubic B-Spline basis functions to represent the contour. For a closed contour composed of \( N \) such elements, the enclosed area may be calculated by:

\[
A = \oint y(s)x'(s)ds = \sum_{e=1}^{N} \int y^e(s)x'^e(s)ds
\]  

(5.1)

The functions and derivative functions can be easily computed given the cubic coefficients defining each element, which may be calculated from the control vertices (see Bartels et al., 1987). The integral in (5.1) is of a quintic equation (arising here from a cubic function multiplied by a quadratic) which, given the curve coefficients, may be calculated exactly by a Newton-Cotes closed formula of degree five requiring five equally spaced abscissas ("Bode's rule", Press et al., (1992)).

Note that areas calculated by (5.1) do not suffer from the biasing implicit in calculating areas from counting a finite number of pixels (see §4.2.2).

5.1.3 Calculation of Area Precision

It is considered more informative to present and discuss results when expressed in the form of fractional precision rather than variance. This is because variance gives no information of the relative size of the variation of area measurements about their mean value in comparison with the size of that mean value, i.e. it is assumed, within reasonable limits, that magnifying a wound image either by altering the imaging geometry or by resampling the captured image gives rise to a proportionate change
in the spread of those measurements. Therefore, variation in area measurements is expressed in terms of fractional precision:

\[
PF = \frac{\sigma_{\text{area}}}{\mu_{\text{area}}} = \frac{\text{S_{area}}}{A}
\]

Note that this equation does not take account of the bias implicit in the estimator for standard deviation. In the following sections the estimates are based on samples of 25 measurements, so that the bias constant for standard deviation estimates is 0.99.

### 5.2 Contour Regularisation Parameters

This section describes the effect of varying the internal energy weights of the GA and TN algorithms upon the mean and variance of the area enclosed by the contour. The same 10 images used for the manual delineation experiments of Chapter 4 were used for this experiment and the 25 manually defined contours produced for each image were used to seed the runs with both algorithms. The image gradient was computed by applying the derivative-of-Gaussian gradient operator to the green-band component of each colour image and used to define the external potential energy given by equation (3.3). It is recognised that the samples of 25 delineations for the ten wounds are composites of biased sets of results (§4.1). Thus the manual precision figure calculated by applying (5.2) to each sample of 25 area measurements contains both intra-delineator variance and inter-delineator variance. Thus (5.2) is a measure of root-mean-square (rms) error. The alternative is to consider bias and precision separately and compare the bias and precision results of each algorithm with the manual biases and precisions. It will be seen in the next section that selecting regularisation parameters for the active contour models described in Chapter 3 is a somewhat complicated issue and the possible combinations of contour regularisation, image regularisation and the need for scale-descent necessitates that a division be made between the two types of regularisation.
The approach taken is to first set the contour regularisation parameters to minimise variance (improve precision) and subsequently to study the effect of scale descent upon both bias and precision.

### 5.2.1 Range and Resolution of Parameter Levels

The combinations of the individual parameter settings used for this experiment, $\alpha$, $\beta$ and the filter scales used to generate gradient potentials at different levels of smoothing, are given by Table 5.1. The external energy weighting parameter was kept constant at 1.0. The choice of parameters tessellates the normalised regularisation parameter plane in a symmetrical pattern, where the normalised parameters may be defined by $\lambda_1 = \alpha/(\alpha + \beta + 1)$, $\lambda_2 = \beta/(\alpha + \beta + 1)$ and $\lambda_3 = 1 - \lambda_1 - \lambda_2$.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Range</th>
<th>Step Size</th>
<th>#Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tension, $\alpha_0$</td>
<td>1/64...64</td>
<td>1 octave</td>
<td>13</td>
</tr>
<tr>
<td>Stiffness, $\beta_0$</td>
<td>1/64...64</td>
<td>1 octave</td>
<td>13</td>
</tr>
<tr>
<td>Gradient Filter Scale, $\sigma$</td>
<td>2...16</td>
<td>½ octave</td>
<td>7</td>
</tr>
</tbody>
</table>

Table 5.1 Specification of parameter settings

Both algorithms were tested on all possible combinations of the three parameters listed in Table 5.1. The analysis is mainly based upon the variance of area measurements, and hence measurement precision, for the following two reasons: (a) the algorithms are generally stable in the absence of external forces and thus are not expected to collapse if the contour is over-regularised, (b) the mean area is expected to vary with filter scale as discussed in §3.1. In the case of the TN algorithm, the stiffness term has a small component that causes the contour to collapse at very high levels of $\beta$. The scale descent procedure is intended to reduce the bias introduced into the area measurement whilst keeping the area variance under control.
The experiments are real world applications of the algorithms to wound images, which are very varied by their nature, thus it is expected that the results of parameter variation will differ from image to image. The results presented here are a representative summary of the general features of the total sum of results obtained.

5.2.2 GA Algorithm: Parameter-Varying Results

Running the algorithm at each level of scale listed in Table 5.1 and calculating the estimated area variance of the 25 contours at each particular combination of \((\alpha, \beta)\) yields a variance map for each level. Instead of observing the variances directly, a more instructive appraisal is obtained by considering the fractional precision of area measurements using the relationship between area variance and precision as given by equation (5.2). The representative features of the precision maps are noted and considered in the following discussion. All precision maps are plotted against the logarithm of the internal parameters, i.e. \(PF (5.2)\) is plotted as a function of \((\log_2(\alpha), \log_2(\beta))\) as shown in Figure 5.1 on the next page.

A. Area Variance Reduced by Increased Image Regularisation

Figure 5.1 shows a representative example of this effect at four levels of image smoothing. As the image smoothing is increased, most images give rise to a variance map with a floor that is relatively smooth and flat over the lower parts of the range of each internal parameter. Only the two most poorly defined images did not exhibit this behaviour. At low scales this floor tends to rise and become less smooth, although at the outer-limits of this floor a small depression may be evident where the variance is optimally reduced by the action of the increasing contour regularisation as evidenced by Figure 5.1(c). The overall effect of variance reduction is due to the image smoothing removing noise and merging together local gradient contours.
B. Variance Increased by Increased Contour Regularisation

The floor of each precision map in Figure 5.1(a)-(c) is surrounded by steeply sloping sides, where the area variance increases towards the level of the manually-defined area variance, as the internal parameters are increased. At all filter-scales tested, no images show a significant deviation from the manual results (i.e. mean and variance) when the internal regularisation is very high. The precision begins to rise when the internal parameters rise above a level which is image-dependent – on some images this is as low as 1 and on some as high as 16. This level also varies to

Figure 5.1 (a)-(d) Precision maps at gradient filter scales $\sigma = 8, 5.7, 4$ and 2.8 pixels respectively.
a lesser extent with filter scale: at higher scales the image forces are weaker and this is manifested by the precision rising at lower levels of the internal parameters. At low filter scales this effect can break down (Figure 5.1(d)) and lead to a band of parameter levels where the variance actually increases above the area variance of the manually defined contours. In this case a few of the contours that define the smallest areas follow the same wound-boundary as the majority of delineations, but at some point deviate inside the wound and come under the influence of a different edge segment. As the internal regularisation is increased through intermediate levels of the internal parameters, the areas of these contours decrease (in contrast to the majority of contours) before increasing again at high levels of the parameters. This large bias in a few measurements causes the variance to increase.

C. Over-Regularising the Image Can Cause Variance to Increase

At the highest filter tested scales two effects become evident on some images that may be regarded as symptoms of over-regularising the image: (a) the floor of the precision map may begin to rise sharply at the very lowest levels of contour regularisation (see Figure 5.2) and (b) the smooth flat floor may start to rise and become increasingly rippled. In the former case the sharp rise in variance at the weakest levels of internal regularisation is caused by the edge of the wound merging with the edge of the limb or another nearby strongly contrasted object as a result of excessive image smoothing. This explanation also partly answers the case for (b). In addition to this the peak attractive forces of an edge are weakened, not only by suppression of
high frequency components of the edge as a result of smoothing, but also by merging of all nearby gradients. In any image of a leg ulcer the background is rarely flat and without contrasted features. These weakened forces compete at different parts of the edge of the wound and attempt to attract parts of a nearby contour.

D. Variance Reduced by Increased Image Regularisation and Decreased Contour Regularisation

The floor of the precision map may sometimes begin to slope as the image regularisation is increased, with the precision gradually improving towards the zero-parameter point \((\alpha = \beta = 0)\) so that the map becomes a continuous rise as the effect of edges merging due to the blurring. Figure 5.3 shows an example of this at two different levels of image regularisation. The effect occurs when epithelialisation tissue is present to a substantial degree at the edge of the wound. This tissue surrounds all or part of the wound and is contrasted both at its interface with the surrounding ‘healthy’ and with the wound and can be ambiguously interpreted by a human observer leading to varied initialisations. At higher levels of image

![Figure 5.3](image_url)  
Continuously decreasing variance at low contour regularisation with no floor  
(a) (left) pronounced effect at high scale, (b) (right) insignificant effect at low scale
smoothing the edges on both sides of the epithelialisation tissue begin to merge with both edges moving towards each other until one edge is created in the middle. This middle point, depending upon the amount of epithelial tissue, can represent a substantial displacement from initial contour positions on both of its sides. Thus as the internal parameters are weakened the contour continuously converge. This explains why the precision map may exhibit a continuous slope at high scales rather than a flattened floor.

5.2.3 TN Algorithm: Parameter-Varying Results

The experiments are repeated in an identical fashion for the TN algorithm using the same initial contours and ranges of parameters. The salient features of the many experimental trials are presented and discussed in this section.

A. Area Variance Decreased by Increased Image Regularisation

The intended effect of image smoothing is to converge multiple contour initialisations to the same position regardless of the algorithm. In contrast to the GA algorithm, however, high levels of internal regularisation do not constrain the final solution to be near to the initial solution. Thus the characteristic ‘floor surrounded by steep sides’ topology of the GA precision map is not shared by this algorithm. Figure 5.4 (a)-(d) shows that the algorithm is generally less sensitive to the setting of the internal parameters than the previous algorithm. In addition to this, the algorithm tends to produce lower levels of variance, being much more able to converge multiple initialisations towards a unique position. This may be observed by comparing the floor of the precision maps in Figure 5.4 to those of Figure 5.3.
B. Over-Regularising the Image Can Cause Variance to Increase

Similar to case 'C' for the GA algorithm (§5.2.2), the TN algorithm becomes sensitive to over-regularisation of the image, especially when the contour is only weakly regularised. Thus similar performance is obtained on images where the edge of the wound merges with a nearby non-wound edge, but only when the image is highly smoothed and contour is weakly smoothed. As stated above, this algorithm has considerable flexibility regarding setting the internal parameters and thus avoiding setting the parameter below a certain threshold is not likely to compromise performance.

Figure 5.4 Example area precision maps for TN algorithm at four levels of filter scale
(a) $\sigma=11.3$, (b) $\sigma=8$, (c) $\sigma=5.7$, (d) $\sigma=4$
C. Excessive Contour Stiffness Can Cause Partial Collapse

Recalling the discussion concerning the contractile effect of the bending forces, it is apparent that when the $\beta$ parameter is set very high, the contour begins to detach itself from the wound outline and may then shrink to an elliptical shape. The shrinkage is caused by the fact that the bending energy is minimised when the contour is a point, similar to the case for the standard elastic energy. However, the high level of $\beta$ also forces the contour to become rather rigid and the forces promoting this rigidity are much stronger in relation to their contractile component (the intended effect of the bending energy term is to impart rigidity to the contour – the shrinkage is more of a side effect). When the contour becomes too rigid, the external forces are overcome by the bending forces and the contour begins to detach itself from the edge of the wound. The effect of such over-regularisation upon the ability of the final contour to represent the wound boundary is dependent upon the shape of the wound. Figure 5.5 (a) and (b) show two example wound images overlaid with final contours obtained with parameters $\alpha=1$, $\beta=64$ and $\sigma=8$. In Figure 5.5(a) the whole contour is partially detached (on the right-hand edge of the wound). If the wound is not particularly elliptical the area can shrink and the result can become a rigid ellipse held in equilibrium by two or three points of high external forces around the contour. An example of this is Figure 5.5(b) where the stiffness of the contour has forced it to detach from both left and right-hand ends and shrink. The effect of excessive contour stiffness on the precision of contour area measurements is observable in Figure 5.4: as $\beta$ increases through the upper-most part of its range, the area precision shows a sharp increase at all levels of scale and at all levels of the elasticity parameter.
Figure 5.5 Examples of the effect of excessive contour stiffness:
(a) Minimal effect on an elliptical-shaped wound, (b) shrinking effect on wound with points of high curvature.

5.2.4 Setting Contour Regularisation Parameters

In order to set explicitly the contour regularisation parameters it is necessary to specify a criterion function that when minimised yields an optimal set of parameters with respect to that criterion. A possible criterion function is to minimise the average precision over the set of ten wounds. This minimisation is then performed for the seven levels of scale parameter $\sigma$ defined in Table 5.1.

\[
(5.3) \quad P_\sigma^* = \min_{\alpha, \beta} \left( \frac{s_{\text{area}}}{A} \right)
\]

This series of calculations is performed for both algorithms and yields somewhat varying choices for different scales. Minimising the average precision over the set of wounds equates to selecting the minimum precision point from an averaged precision map. Instead of selecting explicitly the values $\alpha$ and $\beta$ that completely minimise (5.3) it is possible to manually inspect the averaged precision maps for each algorithm at each scale. Selecting the absolute minimum from (5.3) is more
sensitive to random error than would be the case if a multiple regression analysis were performed (Note: the complexity of the precision maps requires a moderately large number of parameters to model them). The averaged precision maps for the GA algorithm at all scales are topologically very similar, whilst the averaged precision maps for the TN algorithm display increasing levels of noise at low scales. Figure 5.6 shows two example averaged precision maps. In all the maps inspected it has been established that the parameter set \( \alpha = 0.4 \) and \( \beta = 0.4 \) yields averaged precision values that are a close approximation of the absolute minimum values suggested by applying (5.3).

![Figure 5.6 Example averaged precision maps](image)

(a) (left) GA and (b) (right) TN algorithms at an intermediate level of scale (\( \sigma = 8 \))

5.2.5 Summary

*Ground-Attract Algorithm*

Varying both internal parameters appears to have the same effect upon the precision of area measurements, the variances forming smooth concentric contours spread over the parameter plane. The judicious initialisation of the contour near to the
wound boundary represents input from a higher level process, which is maintained as part of the final solution with a weight determined by the contour regularisation parameters. This property is an explanation for the well-behaved precision surfaces in Figure 5.2.

Tangent-Normal Algorithm

This algorithm appears to be more able to reduce area variance than the previous algorithm, although it suffers from the similar problem of losing precision when the image is over-regularised. The flexibility of the contour under a wide range of parameter levels enables contour elements to deform by stretching and translating and, under the control of the stiffness parameter, bend to match any nearby contour. The GA algorithm constrains all three of these degrees of freedom and this may explain why its precision performance is generally poorer.

5.3 Scale-Descent Parameters

The previous section has appraised the general performance in terms of area precision of two algorithms as the contour and image parameters are varied. Precision is naturally improved by blurring the image and the algorithm is stabilised by a small amount of contour smoothing. In this section attention is turned to the subject of recovering the bias introduced by image regularisation, whilst preserving as much as possible of the gain in performance due to improved precision. The mechanism for this recovery is the scale descent algorithm of §3.1. Additionally, the two minimax algorithms, MX and MG, are considered in this section. The objective of this section is to determine a set of parameters for each algorithm that allows optimal performance. The definitions of the performance criteria are considered next.
5.3.1 Performance Measures

Experimental error may be expressed in terms of (a) a function of mean area and (b) a function of variance, corresponding to measures related to bias and precision of area measurements respectively. The view taken here is that the mean manually delineated area is the most representative single quantity defining the area of a wound. A similarity measure to compare the (related) distributions of manual area measurements and corresponding active contour area measurements needs to account for the effect of wound size upon bias and precision measurements. One such measure is simply the square error of the active contour model results with respect to the manual mean area measurement. This measure of comparative performance may be decomposed into a bias and a spread term thus:

\[
MSE = \frac{1}{N} \sum_{i=1}^{N} (a_i - \bar{A}_{man})^2 = (\bar{A}_{acm} - \bar{A}_{man})^2 + s_{acm}^2
\]

where \( a_i \) are the active contour model area measurements.
\( \bar{A}_{acm} \) and \( \bar{A}_{man} \) are the mean areas of the \( N \) measurements for a particular active contour model algorithm and manual delineations respectively.
\( s_{acm}^2 \) is the variance estimate of the algorithm's area measurements.

This measure does not account for the uncertainty invested in \( \bar{A}_{man} \) due to the variance of the manual results themselves. An alternative suggestion is to assign a weight to the error arising between the mean of a set of manual measurements and the associated set of measurements produced by running a scale descent algorithm with one of the four active contour models. The problem is that of combining bias errors and precision errors into one representative quantity, since ultimately a comparative performance measure that is used as a basis for setting some parameter level must be a scalar quantity. Is a bias error less important than a precision error?
A bias between a set of manually delineated areas and a set of areas obtained by running an algorithm can simply be mainly due to the combined uncertainty between the two means. A hypothesis test between the means will serve to disregard all biases below a certain precision-dependent threshold. Consequently, the criterion used here is the ‘safe’ option of using unweighted errors so that equation (5.3) is used as the performance measure for the purposes of setting parameters. This does not however, preclude an assessment of the final results in an in-depth manner.

### 5.3.2 Scale-Descent Results

For each of the four algorithms the scale-descent algorithm is initialised at five scales from \(\sigma=4\) to \(\sigma=16\) in half-octave steps. The scale descent decreases the scale in half-octave steps to a minimum scale of \(\sigma=2\). This process is applied to the ten wound images used for the manual delineation trial. The scale descent algorithm is run 25 times from each starting scale (using the manual delineations) so that the means and variances of measured areas can be estimated. This section discusses the behaviour of the algorithms during scale-descent. The overall behaviours of the four algorithms under the scale descent algorithm of §3.1 are substantially similar. Thus the effects of scale descent can be discussed in a general manner without presenting detailed analyses of each algorithm in turn. The particular differences noted between the algorithms are more associated with the levels of bias and precision errors rather than how the magnitude and ratio of these errors tend to change as the image is de-blurred. This section presents examples cases that are common to all the algorithms. In order to show the variation in both bias and precision and the relationship between them bias is plotted against precision for each scale. Joining the bias/precision co-ordinates obtained at each level of scale produces a locus, which is referred to as the ‘accuracy locus’. For each wound image and each algorithm, five such loci are obtained.
A. Example case of Reduced-Variance with Recoverable Bias by Scale Descent

The image for which the accuracy loci are plotted in Figure 5.7 is an example of a case where the sample manual delineations are mutually biased. This has occurred because epithelial tissue surrounding the partially healed wound has been ambiguously interpreted consistently by one delineator and twice by other delineators. This ambiguity leads to a delineation precision of 6% of the mean delineated area. Applying the GA algorithm with little regularisation of the image fails to substantially reduce this figure because the initial (delineated) contours correspond to more than one energy well and associated minimum of the contour’s energy minimum. Successive increases in initial regularisation of the image introduce increasing biases, but the scale descent procedure allows the biases to be removed, at least to a consistent level of 1-2 %. An increase of this order is expected since the manual mean represents area measurements of more than one distinct wound boundary. The particular examples in Figure 5.7 are taken from the results produced by the GA and MX algorithms.

Figure 5.7  Improved precision after scale descent with recovery of large initial bias
(a) (left) GA Algorithm, (b) (right) MX Algorithm. The dashed line joins the starting points for scale-descents starting at different scales ($\sigma_c$).
B. Example of Permanent Bias Introduced by Over-Regularisation of the Image

Figure 5.6 shows the scale-descent accuracy loci for an image where the loci follow the classical pattern at low levels of regularisation: increased bias at the initial (highest) scale which is reduced to a low level as the scale descent proceeds with the intended effect of providing area measurements with lower variation. This is the case for image regularisation scales up to $\sigma=8$ pixels. However, as the scale is increased beyond this level the loci take on a vastly different form culminating in very large biases, approx. 20%, with no decrease in variation. In this case the image is affected by undesirable illumination effects which, whilst causing no problem of ambiguity to a delineator, causes increasingly large boundary displacement at higher scales. In the example shown in Figure 5.8 the bias is accompanied by a large increase in variance, something which is not typical of the general performance of the algorithms. An additional observation from Figure 5.8 is that the $\sigma=16$ scale descent locus passes close to the (zero bias, zero variance) point at an intermediate level of image smoothing. This is an unusual occurrence of the RMS error being minimised at a high level of scale.

5.3.3 Optimal Scale Parameter Determination

For each algorithm in turn, the RMS error expressing the total error between manual and algorithmic area measurements is calculated for the ten wound images using equation (5.2). Since the error is expressed in percentage terms it is possible to
combine them into a single representative value of the algorithmic performance for each setting of the scale descent algorithm, i.e. start scale and end scale, denoted $\sigma_s$ and $\sigma_e$ respectively. The approach employed here is to take the arithmetic mean of the RMS errors for each wound that arise at each possible pairing of start and end scales, $\Sigma = \{ (\sigma_s, \sigma_e) : \sigma_s \geq \sigma_e \}$. Thus the objective function to be minimised is:

$$E_{RMS}^{\Sigma} = \min_{\Sigma} \left( \frac{1}{W} \sum_{w=1}^{W} E_{w}^{RMS}(\Sigma) \right)$$

where $E_{w}^{RMS}$ is the RMS error for a particular wound.

It is most unlikely that the scale descent pairing yielded by (5.4) will be optimal for each image. Thus, selecting one static pairing of start and end scales will introduce a loss into the overall RMS error, signifying the maximum amount of improvement that could be obtained by an image-dependent scale adjustment procedure. The average loss is defined by:

$$L_{RMS}^{\Sigma} = E_{RMS}^{\Sigma} - \frac{1}{W} \sum_{w=1}^{W} \min_{\Sigma} \left( E_{w}^{RMS}(\Sigma) \right)$$

Table 5.2 contains the results of calculating the RMS error using equation (5.2) for the GA, TN, MX and MG algorithms respectively and lists the start and end scales at which the RMS values are obtained for each algorithm. The average loss computed by (5.5) is also listed and this is subtracted from the overall RMS error value to give an indication of the minimum level of RMS error that could be obtained by selecting appropriate start and end scales for each wound. Figure 5.9 (a)-(d) shows the RMS error and losses for each algorithm on a case-by-case image basis. The darker part of each bar represents the loss for that image introduced by using the fixed scale-descent procedure. These graphs show that the maximum RMS error could be limited to 10% for the GA algorithm, 8% for the MG algorithm and
Table 5.2 RMS errors and scale descent parameters for the four active contour models.

<table>
<thead>
<tr>
<th>Active Contour Model</th>
<th>$E^{\text{RMS}}$ [%]</th>
<th>$\sigma_1,\sigma_2$ [pixels]</th>
<th>$\overline{E}^{\text{RMS}}$ [%]</th>
<th>Predicted $E^{\text{RMS}}$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>4.8</td>
<td>(8,2)</td>
<td>1.0</td>
<td>3.8</td>
</tr>
<tr>
<td>TN</td>
<td>3.9</td>
<td>(8,2)</td>
<td>1.6</td>
<td>2.3</td>
</tr>
<tr>
<td>MX</td>
<td>4.8</td>
<td>(5.7,2)</td>
<td>1.9</td>
<td>2.9</td>
</tr>
<tr>
<td>MG</td>
<td>4.8</td>
<td>(4,2)</td>
<td>1.7</td>
<td>3.1</td>
</tr>
</tbody>
</table>

6% for the TN and MX algorithms. A possible adaptation of the scale selection procedure is to consider the size of the wound. In the present case a good approximation to the size of the wound is provided by the manual delineation procedure. Two issues must be considered when attempting to base the start scale selection upon the wound size:

- If the size of the image smoothing filter used is large and the amount of healthy tissue surrounding the wound is small in comparison with the size of the wound then there is a strong possibility of the wound edge merging with the edge of the limb.

- The size of the smoothing filter required is dependent upon the scale at which the variance is reduced to an acceptable level. If the wound is surrounded by epithelialisation tissue then this scale may well be dependent only upon how much of this tissue is present. This statement anticipates that some delineators may delineate either the inner or outer boundary. Resolving the issue of biased initialisations may require more rigid rules for delineating wounds.
Figure 5.9 Case-by-case analyses of the contribution to overall RMS errors for each active contour model. The total height of each column shows the RMS error associated with each wound image measurement. The darkened part of the bar shows the amount by which the error could be reduced if the optimal scale could be selected specifically for each wound.

5.4 Summary

This chapter has discussed the effect that regularising both the contour and gradient (potential) image has upon the performance of the active contour models described in Chapter 3. The regularisation parameters have been set with regard to precision (contour regularisation) and RMS error (image regularisation). The particular mix of bias and precision that constitutes this error has not been evaluated – this is a subject considered in the following chapter, where the performance of the
algorithms given the specifications for the regularisation parameters is tested upon more images. The MX and MG algorithms do not produce lower levels of RMS error than the GA and TN algorithms indicating that, given the varied set of wounds tested, the exact settings of the contour regularisation parameters are not critical. For the TN algorithm, Figure 5.6 shows that, on average, the precision at all levels of scale employed in the study is quite stable over a wide range of contour regularisation parameters. Due to the nature of its energy formulation the precision of measurements produced by the GA algorithm is progressively degraded as the contour regularisation is increased. There is an unavoidable reduction in the accuracy of measurements produced by all four algorithms as a consequence of choosing working parameters for the scale descent procedure – the ‘loss’ in accuracy, expressed in terms of RMS error, is shown in Figure 5.9.
ACTIVE CONTOUR MODEL EXPERIMENTS

This chapter describes the experiments undertaken to appraise the performance of the four active contour model algorithms described previously in Chapter 3 and incorporates the contour regularisation parameters ($\alpha, \beta$) derived in Chapter 5 for the GA and TN algorithms (recall that the MX and MG minimax-based algorithms are parameter-less with respect to the energy terms) and the scale-descent start and end scales ($\sigma_s, \sigma_e$) for each of the four algorithms. An additional set of ten images is tested with the same parameter and scale settings used for the first set of images and the overall results for the two sets of images are presented. This is considered necessary because the first set of ten images were obtained using sets of optimised parameters determined by the analysis in Chapter 5 and the performance measure is therefore biased in favour of the first ten images – it does not account for the performance of the algorithms upon previously untested images. The aim of this is to produce a more widely based set of results which is not unfairly optimistic due to exclusive use of training-derived data. Secondly, doubling the number of wounds upon which the algorithms are tested from 10 to 20 adds weight to the generality of results and gives a more certain vision of the results that should be expected in clinical practice, which is the ultimate purpose of this analysis.

6.1 Objectives

The active contour models require initialisation by manually delineated wound boundaries forming an initial approximation of the final solution. The algorithms can only be considered a useful tool for wound measurement if it is possible to conclude that area measurements produced by active contour models are an
improvement upon the measurements produced by manual delineation. For the purposes of this analysis, the specific definition of improvement is two-fold and is given by the following statement:

**Definition of Improvement**

To conclude that the application of an algorithm to the wound measurement problem is an improvement upon manual wound measurement it is considered necessary to show:

**Criterion 1**
A significant reduction in the variance of area measurements, taken on a case-by-case basis (i.e. improved precision).

**Criterion 2**
That any differences in area measurements between the manual and algorithmic mean results can be attributed to random variation, rather than being a manifestation of bias.

Therefore, the main objective of this experiment is to determine if the algorithms are capable of producing measurements that conform to these criteria. Thus it is necessary to compare the performance of each algorithm against the manual delineation performance. Complementary to this objective is a comparison between the algorithms in order to determine the levels of measurement bias and precision achievable. Note that testing bias requires comparison with a 'gold standard' of measurement: the approach adopted here is to consider the mean of the manual results, for each wound, to be the best a-priori approximation to such a standard. In practice, the true value of a quantity is defined as the measured value produced by a process that is agreed amongst experts to be an 'exemplar process' (Deming, 1950).
6.2 Experimental Procedure

The four active contour model algorithms described in Chapter 3, viz.: GA, TN, MX and MG, are used to measure the area of each wound image a total of 25 times using each manual delineation attempt for that wound image as the initial contour. The potential energy images are derived from the derivative-of-Gaussian gradient magnitude image of the green band image of each wound image. The GA and TN algorithms both use $\alpha=\beta=0.4$ contour regularisation parameters determined in Chapter 5 and are used in the scale descent algorithm (§3.1), beginning with filter-scale standard deviation $\sigma_s=8$ pixels down to $\sigma_e=2$ pixels. The MX algorithm descends from scale $\sigma_s=5.7$ down to $\sigma_e=2$ and the MG algorithm descends from scale $\sigma_s=4$ down to $\sigma_e=2$. The image filter scale $\sigma$ is reduced in half octave steps.

6.2.1 Additional Image Set

This experiment comprises the original ten images used to select the parameters for the finite element models and ten additional wound images. Of the latter set of ten images (shown in Plate 2), four were taken with a hand-held camcorder with ambient-only lighting (2.1 to 2.4), three were taken with the collimated halogen lighting of the MAVIS instrument (2.8 to 2.10), one was taken with a purpose made illumination rig (2.7) and a further two obtained from a third party wound image library (2.5 and 2.6).

6.2.2 Seeding the Algorithms: Manual Delineation

Two volunteer delineators were asked to provide seed contours for initialising the algorithms in order to run them on the additional image set. In order to make it possible to compare the results with the first set of ten images, for which there were
25 delineations each, the first delineator was asked to make thirteen attempts at delineating each wound under identical conditions to the first manual delineation trial and the second delineator was asked to make twelve such attempts. The wound areas enclosed by the manual boundary delineations are then used to estimate manual precision as a base line for improvement (albeit for only two delineators) by application of the algorithms.

6.3 Development of Analytical Procedures

For each of the 20 wound images the manual delineators produce a set of 25 area measurements. Correspondingly, each algorithm produces a further set of 25 individual contour-area measurements for each of the twenty wounds. A set of area measurements, produced by either manual delineation or by application of an active contour algorithm, may be summarised by calculating the estimates of the corresponding arithmetic mean and standard deviation. These estimates are generated five times for each wound, once for the manual seed contours and four times for the algorithms, providing the data necessary to compare manual and algorithmic performances.

6.3.1 A Note on the Application of Statistical Methods

The results of the first manual delineation experiment in Chapter 4 have been suitably analysed using single-factor fixed effects Analysis of Variance. This requires the area measurements produced at different factor level settings (delineators) to be normally distributed and have equal variances, although these are not absolutely rigid requirements and small departures from them do not much affect the test results. However, the criterion for improvement of precision stated above, requires the variances for area measurements of a wound produced by any one algorithm and manual delineation to be significantly different. The
improvement of precision is the main aim of the application of the active contour models to the problem of wound area measurement.

6.3.2 Improvement of Manual Results

The analytical procedure developed in this section is a comparison between the performance of manual delineation area measurement and the active contour models specified in terms of improvement of precision. The criteria defined in §6.1 may be assessed as a test of relative values and thus conform to standard statistical hypothesis tests. In order to test for variance-reduction (improved precision) the standard F-ratio test for comparing two variances is used. The test employed for bias is Student’s t-test for significant differences between a pair of mean values. The two tests require the definition of the following values:

For any given wound image, let \((\bar{A}_{\text{man}}, s^2_{\text{man}})\) be the respective estimates of the mean and variance of manually delineated area measurements, having mean area \(\mu_\bar{A}\) and area variance \(\sigma^2_{\text{man}}\). Furthermore, let \((\bar{A}_{\text{acm}}, s^2_{\text{acm}})\) be the respective area mean and variance estimates of these quantities produced by application of an active contour. All samples are of size \(N=25\) measurements.

Variance Reduction Test (Criterion 1)

The following equation defines the test to determine if the application of any one active contour model reduces the variance of wound area measurements for a given wound image:

\[
(6.1) \quad \text{Test} \quad H_0 : \sigma^2_{\text{acm}} = \sigma^2_{\text{man}} \\
\text{against} \quad H_1 : \sigma^2_{\text{acm}} < \sigma^2_{\text{man}}
\]
The test statistic is \( F^* = \frac{s_{acm}^2}{s_{man}^2} \). Reject \( H_0 \) if \( F^* < F_{u,v}(\alpha) \), where \( u=v=N-1 \) and the significance level \( \alpha=0.05 \).

**Bias Test (Criterion 2: Equality of Means)**

The test of significant differences between means, i.e. a bias between the mean area measurement of a single wound image produced by a sample of manual delineators and the mean area measurement produced by an active contour model, is defined as:

\[
(6.2) \quad \text{Test} \quad H_0 : \mu_{man} = \mu_{acm} \quad \text{against} \quad H_1 : \mu_{man} \neq \mu_{acm}
\]

The associated \( t \) statistic (modified for unequal variances) is defined as:

\[
(6.3) \quad t = \frac{\bar{A}_{acm} - \bar{A}_{man}}{\sqrt{(s_{acm}^2 + s_{man}^2)/N}} \quad \text{with} \ N-1 \text{ degrees of freedom}
\]

Note: This test is less powerful than the equal-variance \( t \) test which cannot be used here because the standard equal-variance form of expression for \( t \) is not distributed as a \( t \)-variable when the variances are unequal (Johnson and Bhattacharyya, 1992). The consequence of this lower-power test is a broader confidence interval for biases. Also note that this series of tests is not intended to perform a 'multiple-comparison' type procedure among the algorithm performances. Development of a test strategy for comparing the overall precision of the algorithms amongst themselves is considered in §6.3.4.
6.3.3 Development of Graphical Procedure

It is convenient to carry out the tests defined in the last section in a graphical form that allows the test results to be visualised affording easier interpretation. The development of a representative graphical format proceeds as follows:

Firstly, define the axes of the graph by the following equations, for abscissas and ordinates respectively:

\[
X = \frac{s_{acm}}{s_{man}}
\]

(6.4)

\[
Y = \frac{(\bar{A}_{acm} - \bar{A}_{man})}{s_{man}}
\]

(6.5)

These performance measures are plotted for each wound image in Figure 6.1. Equation (6.4) defines a statistic that is related to the F-ratio test of (6.1), thus the abscissa value of a point on the graph may be used to test for the reduction of variance. From standard tables of the F-distribution the critical value of $F_{24,24}$ at a significance level of $\alpha=0.05$ is 0.5. This allows the definition of a confidence interval for the variance of measurements made by manual delineations. The result is a pair of vertical lines on the graph at $X_{lower} = 1/\sqrt{2}$ and $X_{upper} = \sqrt{2}$ signifying the limits for improved variance and degraded variance respectively.

Equation (6.5) defines a test statistic that is related to the t-test of (6.2), allowing the equality of the mean area measurements of a wound, produced by manual delineation and an algorithm, to be evaluated according to the vertical position of a point on the graph. Let $t_{critical}$ be the critical value of $t$ with $N-1$ degrees of freedom (df) for a two-sided test at a significance level of $\alpha=0.05$. Equation (6.3) may now be re-written in terms of (6.4) and (6.5) as:
From tables of the 2-sided $t$ distribution with 24 df, $t_{critical}$ has a value of 2.064. Thus in this application (6.6) becomes:

$$Y_{conf} = \pm t_{critical} \sqrt{\frac{1 + X^2}{N}}$$

Again, (6.7) defines a pair of upper and lower confidence intervals that may be plotted on the graph.

The area measurement results in terms of measured means and variances are expressed in terms of the manual standard deviation so that an improvement in precision is clearly shown, without regard to the actual levels of manual and algorithmic fractional precision. Therefore, this diagram allows for the drawing of two confidence intervals, one for variance and the other for bias which are equally valid for all of the wounds measured so that the significance of all wounds may be tested and displayed on the same diagram (see Figure 6.1).

### 6.3.4 Comparison of Algorithms: Overall Test for Precision

The previous analysis allows the comparison of each algorithm's performance with the manual delineation performance on an image-by-image basis. However, this does not necessarily allow a general comparison to be made between the algorithms over the whole 20 wounds. Since the variances are intended to be altered significantly by the application of each algorithm, and the variances will inevitably be image dependent, a non-parametric or rank test using blocking to remove the variation amongst wounds is used. The Friedman test meets such criteria. The basic
datum for the test in this case will be the range of measurements produced by each algorithm upon each wound. The test is defined as follows:

\[
(6.8) \quad r_{w,a} = \max_i (A_{w,a,i}) - \min_i (A_{w,a,i}) = E_w + R_a + \varepsilon_{w,a}
\]

where \( E_w \) is defined as the range block-effect that is a random variable dependent upon the wound properties and its size. Note from Chapter 4 that both of these quantities affect precision. \( R_a \) defines the general range (precision) of each algorithm. Note that \( w \in \{1..W\} \) and \( a \in \{1..k\} \), where \( k=5 \). The 'treatment' groups denoted by 'a' correspond to the four algorithms and manual delineation, giving a total of \( k=5 \) groups. \( A_{w,a,i} \) corresponds to the \( i \)th area measurement of wound \( w \) made with algorithm \( a \), where \( i \in \{1..25\} \).

The null hypothesis states that the general range or spread of the area measurements, and hence the true variation of each algorithm’s performance is the same. It is defined as follows:

\[
(6.9) \quad H_0 : \text{All } R_a \text{ are equal}
\]
\[
(6.9) \quad H_1 : \text{Not all } R_a \text{ are equal}
\]

For a significant difference in range to be detected an algorithm must consistently perform better than the other algorithms. If \( H_0 \) is rejected, and therefore the algorithms have definite differences in overall precision, it becomes necessary to consider which algorithms differ. Dunn’s multiple paired-comparison procedure is suitable for this purpose (see Neave and Worthington, 1988). The paired differences computed within each block (wound) are approximately distributed as \( N(0, Wk(k+1)/6) \). This allows the construction of Gaussian confidence limits for the rank sums from the Friedman test, given by:

\[
(6.10) \quad \overline{R}_a \pm \frac{1}{2} z_\alpha \sigma \quad \text{where } \sigma = \sqrt{Wk(k+1)/6}
\]
When the confidence limits are plotted on a line diagram, overlapping ends imply that the ranges of area measurements produced by the two corresponding processes do not differ significantly at the $\alpha$ significance level. Note that the confidence limits apply to the rank sums and not to the original range data.

### 6.3.5 Quantification of Bias and Precision

The previous sections define the analytical method for the improvement of manually delineated wound measurements without regard to percentage values of precision or bias. The errors are analysed in terms of variance ratios (precision related) and systematic error (bias related). By their very nature, such relative values and their related hypothesis tests do not contain information about the size of either a precision or bias error. A significant bias in measured area is strong evidence that an active contour model is in general delineating a different boundary to the corresponding mean manual boundary, by including areas not defined as wound by the manual operators or by excluding areas which are defined as wound. However, when comparing manual and algorithmic measurements, obtaining a significant value from an appropriately applied hypothesis test does not necessarily indicate that the bias difference represents a large portion of the mean wound area. For wounds where the precision of both sets of measurements is relatively high, a hypothesis test such as the $t$-test will be correspondingly more sensitive to bias errors, i.e. the likelihood of detecting smaller fractional biases will increase. The absolute value of this bias is important, since when measurement precision is high, a relatively small bias is a small fractional error which introduces little variation in the final result despite being a major part of the overall error. Clearly, a large fractional bias that is significant can never be dismissed in such a manner. In the cases of wounds where the precision of measurements is poor, the discrimination of bias from the expected mean error under the equal-means hypothesis becomes increasingly blurred. This section therefore proposes an additional method of analysis that accompanies the previous one and which will allow the definition of
the performance of the algorithms defined in this thesis to be reported explicitly in terms of the average percentage bias and precision.

**Bias Measurement Method**

The bias performance of each algorithm is assessed by calculating the mean of the absolute biases that exist between the manual and respective algorithmic measurements taken from each wound in this study. The mean absolute bias is calculated as follows:

Firstly, the individual bias for each wound, \( w \), is defined as:

\[
B_w = \frac{\bar{A}_{a} - \bar{A}_{m}}{\bar{A}_{m}} = \frac{\bar{A}_{a}}{\bar{A}_{m}} - 1
\]

(6.11)

The estimate of the standard error is dependent upon the ratio \( \frac{\bar{A}_{a}}{\bar{A}_{m}} \) in (6.11) and can be shown to be:

\[
s\{B_w\} = \frac{1}{\bar{A}_{m}} \sqrt{\frac{1}{N} \left( \frac{s_{a}^2}{\bar{A}_{a}^2} + \frac{s_{m}^2}{\bar{A}_{m}^2} \right)}
\]

(6.12)

The overall mean absolute bias for the set of wounds is then defined as:

\[
MAB = \frac{1}{W} \sum_{w=1}^{W} |B_w|
\]

(6.13)

where \( W \) is the number of wounds over which the average is taken.
This expression has an estimated standard error of estimate given by:

\[
(6.14) \quad s\{MAB\} = \frac{1}{W} \sqrt{\sum_{w=1}^{W} s^2\{B_w\}}
\]

**Precision Measurement Method**

Complementary to the bias calculation is the quantification of the average level of precision achieved by each algorithm. The mean fractional precision for each algorithm is defined by the following equations:

The fractional precision for a set of area measurements of one wound is defined as:

\[
(6.15) \quad PF_w = \frac{s_w}{A_w}
\]

The estimates of fractional precision are made with samples of $N=25$ measurements, thus the omission of the correction factor $c_4$ associated with estimates of standard deviation has little influence on the results. The standard error of (6.15) may be estimated from:

\[
(6.16) \quad s\{PF_w\} \approx \sqrt{\frac{s^2_w/(N-1)}{A_w^2}}
\]

where it is noted that the standard error of a standard deviation estimate based on a sample of size $N$ is closely approximated by $\sqrt{s^2/(N-1)}$. Taking the mean precision for a set of wounds (sample size $W$) yields the following estimator for mean fractional precision and its estimator of the associated standard error:
\begin{equation}
PF_{av} = \frac{1}{W} \sum_{w=1}^{W} PF_w
\end{equation}

\begin{equation}
s\{PF_{av}\} = \frac{1}{W} \sqrt{\sum_{w=1}^{W} s^2 \{PF_w\}}
\end{equation}

The manual mean area for a wound, $\bar{A}_{man}$, is used as the reference point for all measurements, since it represents the combined opinions of several manual delineators. Although it is known that delineators are in general mutually biased (§4.1.7), considering any one subset of the measurements to be the best estimate of a wound's area could be judged to be introducing a personal bias into the results.

### 6.4 Experimental Results

Appendix C contains tables of the bias and precision estimates of the wound area measurements arising from application of the algorithms to the twenty wound images. Estimates of absolute bias and precision for the application of each algorithm to each are calculated using (6.11) and (6.15), along with the associated standard errors, calculated with (6.12) and (6.16) respectively. The data in these tables is used to plot the graphs presented in this section.

#### 6.4.1 Detection and Analysis of Outliers

The measurement of the spread or variation of area measurements is highly sensitive to the presence of outliers in the data. Thus, the hypothesis tests for variance reduction (6.1) and bias (6.2) along with the calculation for precision (6.15) and most definitely the calculation for range (6.8) will be affected. There are three possible sources of outliers in the area measurements:
The manual delineation data for the first set of images has been analysed for outliers in Chapter 4. There are four area measurements considered as outliers, however, these measurements do not correspond to faults and so are not discarded. The 25 area measurements made by each algorithm are checked for outliers, for each wound, with the following rule:

Label \( A_i \) as an outlier if 
\[
A_i \leq Q_L - \frac{3}{2}(Q_U - Q_L) \quad \text{or} \quad Q_U + \frac{3}{2}(Q_U - Q_L) \geq A_i
\]

Where \( Q_L \) and \( Q_U \) are the lower and upper quartiles of each measurement sample respectively.

Applying this rule to the basic area data tends to indicate one or more outliers in approx. half the 20 wounds cases for each algorithm, although no case can be attributed to a procedural or equipment fault. Inspection of the regions delineated by each algorithm shows that outliers tend to be indicated when the data is divided into two groups, one being larger than the other, e.g. 20 measurements in one group and the remaining 5 measurements forming a distinct second group. This occurs because of the ambiguity of the edge evidence present in the images. Since the algorithms first seek equilibrium at an increased initial scale the areas of the delineated regions produced by the algorithms at the initial scales may be similarly inspected for outliers. The result of this inspection shows that few cases of outliers are now indicated. Thus, at high scales, where the 'multiple edges' are merged, the data tend to be clustered (except for the most ambiguous wound cases). As the scale is reduced, the merged edges begin to break up and the noise also increases. This
combination exploits the small differences that exist in the boundaries produced at higher scales and causes them to diverge as the scale is lowered.

6.4.2 Relative Improvement of Algorithms

Figures 6.1 (a)-(d) show the combined results of the hypothesis tests for bias and improved precision defined by (6.1) and (6.2). These diagrams show the results of the algorithm trials in relative terms only, i.e. the diagrams show whether or not the variance of a set of area measurements made by an active contour model on a particular wound is a significant improvement upon the corresponding manual variance and also show if the measured bias differences are insignificant. There is no information in these diagrams to show fractional (percentage) errors, e.g. a point at (0.5,1.0) indicates that the standard deviation of a measurement sample produced by one of the algorithms for a particular wound was half of the corresponding manual standard deviation, regardless of whether the manual standard deviation equates to a precision level of 1% or 10%.

Likewise, the abscissa value of 1.0 represents a statistically significant bias but does not indicate the magnitude of bias expressed as a percentage of the mean manual area. If both manual and algorithmic measurements of a wound's area give rise to small values for precision, then only a small bias is needed to signal a significant difference. If there is a spreading of either the manual or algorithmic precision then a corresponding increase in measured bias is required for a given significance, else the significance decreases and any true bias is swamped by the variation in results. The confidence limits in Figures 6.1 (a)-(d) that define the regions corresponding to the outcome of the variance reduction test (6.1) are plotted using (6.3). Equation (6.6) defines the function used to plot the upper and lower mean-difference confidence limits. The combination of these two separate intervals defines a key region on the diagram: this region is enclosed between the two \( t \) confidence curves and bounded on its right side by the lower \( F \) confidence limit. Marked points that
Chapter 6  Active Contour Model Experiments

(a) GA Algorithm

(b) TN Algorithm

(c) MX Algorithm

(d) MG Algorithm

Figure 6.1 Improvement Diagrams

The filled circles represent measurements of wounds from the initial (training) set. The squares represent measurements made on the additional set of images. Image 2.10 is out of range in all graphs due to a large negative \( r \) value. Image 2.2 is out of range on graphs (a), (c) and (d) for the same reason. Likewise, image 2.8 is out of range on graphs (b), (c) and (d) and image 2.6 is out of range on graph (b).

lie within this region represent algorithm measurements of wounds where there is a significant reduction in variance and no statistically discernible bias.
The central point at (1.0,0.0) in Figures 6.1 (a)-(d) represents the standard deviation and mean estimates for the manual delineation results for any wound. As a basic summary of the performance of each algorithm with respect to the manual delineation performance, Table 6.1 displays a count of the number of times each algorithm made the following specific improvements:

(a) Rejected the equal-variance hypothesis $F$-test given by (6.1), i.e. produced significantly lower variance for a set of wound measurements (Criterion 1).

(b) Did not reject a $t$-test of mean differences given by (6.2), i.e. the algorithm is not considered to have introduced a bias (Criterion 2).

(c) Both above cases were coincident and thus the point lies in the 'key region' mentioned above, i.e. complete improvement (both Criteria).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Criterion 1 Significant improvement</th>
<th>Criterion 2 Equation 6.2 Not rejected</th>
<th>Both Criteria Met</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>14</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>TN</td>
<td>18</td>
<td>12</td>
<td>12</td>
</tr>
<tr>
<td>MX</td>
<td>14</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>MG</td>
<td>14</td>
<td>8</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 6.1 Frequency of occurrences of agreement and precision improvement criteria

The hypothesis test outcomes for the application of the four algorithms to individual wounds is presented in a side-by-side comparative form in Table 6.2. From this table it can be observed which wound measurements were improved by most or all of the algorithms and which wounds tended to lead to poor measurements by the algorithms.
166

6.4.3 Overall Precision Test Results

The result of the hypothesis test defined by (6.9) defined for testing the equality of the general precision of the algorithms and the manual delineations is summarised in Table 6.3. The confidence intervals for the rank sums defined by (6.10) are used to produce the line plots shown in Figure 6.2. The overall grouping of the different algorithms is clear from this diagram. The results are discussed in Chapter 7.
### Chapter 6  
Active Contour Model Experiments  

#### 6.4.4 Bias and Precision Results

The basic measurement data contained in Appendix C, which was presented in the foregoing section in terms of improvement upon the manual performance, is presented here a second time but in a different guise. As stated previously, the improvement data is expressed in terms of the manual precision which may be directly interpreted in terms of a hypothesis test but which contains no information about the sizes of random and systematic errors in relation to the 'true' area of a wound. The general performance of each algorithm in terms of bias is presented in Table 6.4 as mean absolute bias accompanied by the counts from Table 6.1 for the number of times the $t$-test for mean differences was not rejected for each algorithm. Mean absolute bias and its standard error is estimated with (6.13) and (6.14) respectively. Average fractional precision is estimated with (6.17) and its standard error is estimated with (6.18).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>$MAB \pm \text{std. error}$ [%]</th>
<th>$PF_{ev} \pm \text{std. error}$ [%]</th>
<th>#(6.2) not rejected</th>
<th>#(6.1) rejected</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>3.4 ± 0.2</td>
<td>2.4 ± 0.2</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>TN</td>
<td>4.0 ± 0.2</td>
<td>1.8 ± 0.2</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>MX</td>
<td>3.8 ± 0.2</td>
<td>2.6 ± 0.3</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>MG</td>
<td>2.8 ± 0.2</td>
<td>3.1 ± 0.3</td>
<td>8</td>
<td>14</td>
</tr>
<tr>
<td>Manual</td>
<td>NA</td>
<td>4.3 ± 0.3</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 6.4 Mean absolute bias and mean precision for algorithms
Chapter 6  Active Contour Model Experiments  168

(a) GA Algorithm

(b) TN Algorithm

(c) MX Algorithm

(d) MG Algorithm

Figure 6.2  Bias-precision plots

The filled circles correspond to measurements made on the first set of images. Although the diagrams do not indicate the significance of either of the hypothesis tests previously described, the practically significant cases where bias is large are evident.

The bias and precision estimates for individual wounds are displayed in Figures 6.2 (a)-(d) for the four algorithms. The dashed rectangle plotted in each diagram bounds the region where bias is within ±5% and fractional precision < 5% and is included purely as a visual guide for the reader.
6.5 Summary

This chapter has presented the results of running the four active contour models described in Chapter 3 to twenty images of leg ulcers. The first objective of the presentation has been to display the performance measures defined in §6.1 that compare the observed performance of the algorithms with the observed performance of the manual delineators (results from §4.1). The performance measures are intended to indicate (a) whether the algorithms can reliably reduce the variability inherent in manual delineation and (b) whether the results of algorithmic area measurements agree with those produced by manual delineation. Secondly, a performance measure has been devised to compare the average measurement precision produced by the algorithms. Emphasis has been placed on the need to evaluate the performance of the algorithms over a varied range of wound images, since such wounds are likely to frequently occur in practice. The results presented will be discussed in detail in the next chapter, where particular attention will be paid to those wound images where the algorithms fail to produce measurable improvements.
7 DISCUSSION

7.1 General Performance of the Algorithms

Precision

Table 6.1 shows that all four algorithms produce significant improvements in the variance (precision) of area measurements in most of the test cases. The comparison for average precision between the algorithms given by Table 6.3 shows that there are detectable differences in precision between the four algorithms. The differences between the algorithms and the manual precision (including inter-delineator biases) are clearly shown in Figure 6.2: the TN algorithm shows the greatest measurable improvement, although its confidence interval overlaps to a small extent with that of the MX algorithm. The general precision performances of the GA, MX and MG algorithms are not clearly separable, and it should be noted that there is a small overlap between the manual confidence interval and that of the MG algorithm. More information about the performance of the algorithms in terms of precision may be gleaned by considering the number of improvements in precision made. The TN algorithm produces 50% more cases of measurable improvements in precision of the manual results than do the other three algorithms, the algorithm showing 18 out of 20 cases of improved precision, versus 14 out of 20 for each of the other algorithms. Conversely, the TN algorithm fails to produce a measurable improvement of precision in two cases, and the other three algorithms each fail to produce measurable precision improvements in six cases. This increase in the number of improvement cases is thus in agreement with the results of the overall precision test.
Bias

In contrast to the case for precision, the four algorithms experience some difficulty in returning non-significantly biased results, with no algorithm able to return non-significant differences for more than 12 of the wounds. In this respect the TN algorithm again produces the highest number of successful cases, showing significant bias in only 8 out of 20 cases, in comparison with the other three algorithms which show significant biases in 12 out of 20 cases. Since the TN algorithm also produces the best precision performance, the decreased incidence in the cases of bias for this algorithm cannot be attributed to higher variances causing less significance to be shown in the t-tests.

Combined Improvements

The number of times both performance criteria are met simultaneously again highlights the differences in the overall performance levels of the algorithms. The TN algorithm produces 12 such cases; i.e. showing significant increases in precision and having non-significant differences in mean areas. In this respect the other algorithms perform much more weakly, with the minimax algorithms showing simultaneous improvements on less than six wounds. From Figure 6.1 it is clear that the major contributor to the failure rate is bias. Since the highest simultaneous success rate is only 60% (12 out of 20 wounds) for algorithms tested here, a large part of this discussion is devoted to finding the causes of the failures.

7.1.1 Precision Measurement

Table 6.4 shows that the mean precision for manual measurements of 4.3±0.3% (25 measurements) is reduced by each algorithm. The best average precision of 1.8±0.2% is achieved by the TN algorithm – this is in agreement with the overall precision test and the number of precision improvement cases stated above. The
lowest estimated precision is produced by the MG algorithm (3.1±0.3%). The GA, MX and MG algorithms each show one instance of significantly increased variance along with additional cases of non-significant increases that are discussed later in this chapter. In 19 out of 20 cases the highest measurement precision is achieved by either the TN algorithm (14 cases) or the GA algorithm (6 cases), there being one case of a tie between the two algorithms. The GA algorithm shows significant increases in precision in comparison to the TN algorithm for images 1.6 and 2.3; these cases are examined here:

A Image 1.6

Both algorithms produce the same bias value (+2.0%) for this wound and start and finish scale-descent at the same scales, but the GA algorithm has a precision of 1.2% and the TN algorithm, 3.4%. Figures 7.1 (a) and (b) show a comparison of the median boundaries of the delineations produced by the GA and TN algorithms. Note that the boundary produced by the TN algorithm is substantially smoother than that produced by the GA algorithm. The boundary of the GA algorithm is a closer match to the wound's edge, deviating mainly from the manual boundary by missing a slightly elongated portion at the top of the wound and having a slightly thicker 'tail' at the right-hand side. It is clear that the very nature of the GA algorithm has enabled it to supersede the performance of the other algorithms in this instance.

B. Image 2.3

Figures 7.1 (c) and (d) show a comparison of the median boundaries produced by the GA and TN algorithms. A situation similar to the one described above arises here: the GA boundary is much less smooth than the TN boundary and is clearly a better delineation of the wound's boundary, just failing to converge fully with the upper-left hand edge of the wound. The failure of the TN algorithm in this case is due to the effect of high image smoothing upon this particular wound because of its shape. Note from Table 6.2 that the measurements of all algorithms met both
performance criteria stipulated in §6.1. Thus, even though the precision of the TN algorithm is poorer than that of the GA algorithm, it is still a definite improvement upon the manual precision. Additionally, despite the fact that the algorithm clearly fails to properly delineate the wound, the bias difference is insignificant. Despite these facts, the 'failed' delineation is visible and such a result in practice can therefore be treated, appropriately, with caution.

Figure 7.1 Comparison of median boundaries representing average wound delineations
(a) and (c) the GA algorithm and (b) and (d) the TN algorithm.
It should be noted that the action of the stiffness parameter $\beta$ in the TN algorithm is to promote smoother boundaries, whereas in the GA algorithm all second order deviations from the initial contour cause a progressive increase in contour energy, without regard to whether the contour is becoming more smooth or less smooth. The contour energy of the GA algorithm will increase if the action of the image gradient forces is to smooth the initial boundary. This will occur when applying a manual seed contour at a high scale, and is an explanation of why the GA algorithm does not converge with such a high precision as the TN algorithm. The two cases described above are examples of a situation where the gradient energy is insufficient to overcome the smoothing forces in the TN algorithm, i.e. for these image cases, the contour is over-regularised.

The MG algorithm is the weakest performing algorithm in terms of precision. It can be seen from Figure 6.1(d) that 9 of the significant improvements in precision lie in a narrow vertical band between $X=0.5$ and $X=0.7$. The cases of improved precision of the other algorithms tend to be scattered closer to the $X=0$ abscissa. Also, it is clear that this algorithm does not create any more cases of variance increases than either the GA or MX algorithms. The band may be explained thus: The contours of minimum energy for the two external forces, the gradient forces and gray-level forces, do not tend to coincide with each other, so that convergence with one contour means that the energy term for the other external energy will not be locally minimised, and vice-versa. This competition between the two external energies is a contributory factor in producing variance in the results and may therefore be used as an explanation of the higher levels of variance experienced with this algorithm.

### 7.1.2 Bias Measurement

The lowest average absolute bias value of $2.8 \pm 0.2 \%$ is produced by the MG algorithm (Table 6.4), although this algorithm shows 12 cases of significant bias. In contrast, the TN algorithm, which shows just 8 cases of significant bias, produces
the highest absolute bias average of 4.0±0.2%. Figure 6.2(b) shows that the TN algorithm produces three cases of large bias, two in excess of 10% and one in excess of 20%, all of which are significant (images 1.10, 2.8 and 2.10). In comparison, Figure 6.2(d) shows that, although the MG algorithm also produces its largest biases on images 2.8 and 2.10 it contained these biases at the 8-9% level, and it does not produce a significant bias on image 1.10 – the bias is less than 5%. The large bias is a starting scale problem: Figure 7.2 shows some example median boundary delineations for image 2.10 that represent the average of the converged results from the TN algorithm and the MG algorithm. It is clear from these images that the problem is one of over-regularising the image: both algorithms fail to delineate the ‘tails’ of this wound. In fact it is evident that the MG algorithm actually performs worse than the TN algorithm when both algorithms are started at a high scale. This suggests the need for a shape-based control to adjust the starting scale. This information could be obtained directly from analysis of the initial contour or alternatively, the person delineating the wound could be presented with several shape templates and asked to indicate which one best approximates the wound. Although the TN algorithm produces fewer cases of significant bias, its performance can vary significantly on some wounds. These cases are discussed in the next section.

![Figure 7.2 Comparison of median boundaries for TN and MG algorithms applied to image 2.10.](image)
7.2 Analysis of Failure Cases by Wound

A wound measurement made by one of the active contour algorithms may be labelled a 'failure case' when either one or both of the following failure modes apply:

- Its precision degrades with application of the algorithm.
- The algorithm's mean measurement is significantly different from the manual one.

The purpose of this discussion is to identify a set of properties in the wound images that explains the errors. Both of the failure modes occur a total of three times among the measurements, made on wounds 2.8 and 2.10 (see Table 6.2). These cases are also the only ones where variance significantly increased. Note that the variance ratio test for identifying cases of improved precision produces three groups of results, (a) significantly decreased, (b) not significantly different (ratio of variance estimates may be numerically less than or greater than 1.0 but not enough to be considered significant) and (c) significantly increased. The improvement criteria stated at the beginning of this chapter include case (a) whereas the failure mode for variance test is case (c) – this ignores wounds that are cases of (b).

7.2.1 Cases of Both Failure Modes

Two of the three cases where application of an algorithm caused a degradation in precision and a significant bias refer to wound 2.8 with the GA and MG algorithms. This image is of a sloughy (yellow) wound on the upper side of the foot just above the toes. Subjectively, the wound appears distinct from the surrounding skin, having a different hue and regular shape (ovoid). In this case the reflectance of the wound tissue is actually higher than that of the surrounding skin and inspection of the edge of the wound shows colour changes – a small amount of red epithelialising tissue,
purple-hued swollen tissue and some white flakes of skin – which produce multiple edges. Also, there is a band of shadowing by the toes which at low filter scales will produce two edges having opposing directions.

The MX algorithm gives rise to a significant bias and a significantly increased variance for wound 2.10. The bias differences for this wound are the largest recorded for the set of 20 wounds, ranging from 8.6% to 21.7%. Figures 7.2(a) above, show the general error that has occurred in this image: the final contour has pulled away from two of the three points of highest curvature which have little support in the gradient magnitude image at the starting scale. There is also a red dark spot (probably pooled blood, as opposed to blood perfusing through healing tissue) at the neck of one of the tails of the wound which gives rise to high gradients at its periphery. The fact that two of the algorithms (TN and MG) produce significantly improved precision values for this wound (both 1.2%) is of no real consequence due to the dominance of bias.

7.2.2 Insignificant Cases of Variance Ratio

Table 6.2 shows that images 1.9 and 1.10 returned increases in variance that were not statistically significant. Both of these wounds are examples of some of the most ambiguous images in the wound library. Table 4.1 in Chapter 4 shows that all five manual delineators perform at their least consistent with these two images, returning average biases of 18% and 26% respectively. It should also be noted that there is little consistent gradient information in these images. Thus, in this case there is insufficient image evidence to allow for these differences of opinion to be converged together. The capacity for variance-reduction with image 2.4 is limited because of the vagueness of the boundary at the top of the wound that prevents all of the algorithms from converging to a repeatable result.
7.2.3 Cases Where Bias is Significant

As stated earlier in this chapter the main manifestation of error in the results is the number of cases where most or all of the four algorithms tested produce significantly biased results when compared with the manual results. Notably, there are five wounds in the set of 20 where all four algorithms introduce a statistically significant bias (see Table 6.2) and there are a further six cases where three of the four algorithms produce significantly biased results. In four of these latter six cases, the TN algorithm is the only one that produces insignificant biases. Whilst these biases are deemed significant, the size of the bias is still an issue that requires addressing. The following table summarises the failures.

<table>
<thead>
<tr>
<th>Image</th>
<th>Summary</th>
<th>Bias (% mean area)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>GA</td>
</tr>
<tr>
<td>1.3</td>
<td>All algorithms fail</td>
<td>3.4</td>
</tr>
<tr>
<td>2.2</td>
<td></td>
<td>-4.1</td>
</tr>
<tr>
<td>2.8</td>
<td></td>
<td>4.3</td>
</tr>
<tr>
<td>2.9</td>
<td></td>
<td>-3.6</td>
</tr>
<tr>
<td>2.10</td>
<td></td>
<td>-14.0</td>
</tr>
<tr>
<td>1.1</td>
<td>3 algorithms fail</td>
<td>-3.9</td>
</tr>
<tr>
<td>1.4</td>
<td></td>
<td>3.0</td>
</tr>
<tr>
<td>2.1</td>
<td></td>
<td>-1.7</td>
</tr>
<tr>
<td>2.4</td>
<td></td>
<td>4.9</td>
</tr>
<tr>
<td>2.5</td>
<td></td>
<td>-1.2</td>
</tr>
<tr>
<td>2.6</td>
<td></td>
<td>0.4</td>
</tr>
</tbody>
</table>

Table 7.1 Cases of wounds that give rise to significant bias for most algorithms.
Values marked in bold typeface are not statistically significant.

Some biases are statistically significant whilst being relatively small, e.g. 1.7 % bias recorded for image 2.1, and represent the cases where an algorithm tends to miss a small part of the wound, rather than a catastrophic failure. It is not clear whether such bias errors can be systematically propagated through a series of
wound measurement samples taken at regular intervals over a period of a wound’s occurrence, since a wound can be expected to change its appearance, particularly if it is changing in size. Clearly though, when an algorithm fails to produce a ‘correct’ measurement (e.g. images 2.8/2.10) and the bias is definitely not a difference of opinion between the edge evidence and the manual delineator, it renders the measurement useless. Since the error is so large, it allows easy visual detection by the delineator so that the algorithm measurement can be discarded and the manual result maintained as the best estimate of the wound’s area. Small errors are not necessarily obvious, since one expects a small ‘noisy’ variation in the results, and in a clinical/work-place setting a delineator may not be able to make a series of measurements of a wound.

7.2.4 Identifiable Causes of Bias Errors

By examining the median boundary of each wound measurement sample, three features present in the wound images tested have been identified to which systematic error is attributable. Some of the causes have been identified above in the discussion of failure modes.

(1) Vague boundaries

The bias errors on images 2.2, 2.4 and 2.6 are attributed to lack of clear image edge evidence. This occurs at the lower edge of 2.2 (see PLATE 2), the top part of 2.4 and most of the right hand side of 2.6. The biases are all in the 2% to 4% range.

(2) Redness at the edge of the wound

Images 1.3 and 2.8 produce maximum biases of 3.5% and -17.7% respectively. In the case of 2.8, precision is also affected. The redness is
epithelialisation tissue highly perfused with blood which turns a purple hue after a period of time (as the wound heals). This tends to produce multiple boundaries and the algorithms therefore tend to produce results that include parts of both boundaries.

(3) Darkened spots near the boundary

Cases of this are wound images 1.1 and 2.10. In the case of 1.1 the bias error is limited to 3.9% and the effect of the spot (blood inside the wound) is to provide an alternative high-gradient boundary segment which attracts the contour.

7.3 Consequences of Errors

The ability of the algorithms to significantly improve area measurement precision has a practical purpose – it gives the physician a higher degree of confidence that a sample area measurement is close to some mean value. The disadvantage is that this mean value can, and often is, biased with respect to the average human judgment. This does pose a problem for the use of the algorithms in clinical practice: the objective in measuring wounds is to compare measurements taken at regular times separated by intervals of possible change to the wound's size and appearance. A wound may be classified as 'indolent', indicating no change in size or colour, or it may be healing (getting smaller) although this may be due more to volume changes rather than area changes in which case area measurement is less informative. Alternatively a wound may be infected, which may cause it to change colour and probably increase in size.

The results presented here represent a medium sized sample of the wounds that can be encountered in clinical practice. In a hospital outpatients clinic it is not always possible to take many measurements of a wound and perform a hypothesis test to
determine whether a wound has changed size. In the event of a hypothesis test between two measurement samples, a significant difference can be interpreted as a genuine change in a wound's area, conclusive evidence of healing or deterioration (perhaps due to infection), or it may be due to a systematic error (bias) between the two samples.

Detection of wound size changes can be improved by taking a set of measurements for the purposes of performing a hypothesis test where the null hypothesis is that the mean values for a current wound measurement is the same as the one previously taken. The equal variance $t$-test for differences between means when the sample sizes are equal, is given by (Johnson and Bhattacharyya, 1992):

$$t = \frac{\overline{A_1} - \overline{A_2}}{s_{pool} \sqrt{2/N}}$$

with $2(N-1)$ df.

Comparing two measurement samples of five measurements each, a mean difference of approx. $1.5s_{pool}$ is required for significance. If the delineator's precision ($s_{pool}$) is 5%, then a mean difference of 7.5% is required between the sample means for a change in the wound's area to be signalled. The sensitivity can be increased by increasing the sample size, although this requires more time to be made available for delineation and it is possible that a delineator's performance may degrade as a result. The precision may be improved by applying an active contour model, thus alleviating some of this burden by allowing a smaller sample size to be used without the test losing its power to detect a change in wound size. However, one must also consider the effect of the bias errors introduced by the algorithm upon this difference. Considering the above causes of bias, it is possible that such errors will tend to be either all positive or all negative given a sequence of measurements taken over some period of time. Thus, the bias error may be a fairly constant component of the measurements thus having a relatively small effect on the hypothesis test. However, when a bias error is very large and the resulting contour has not managed to properly delineate a fairly clear wound boundary (e.g. image 2.10) the error
should not be ignored. As previously stated, errors of this type are visible to the
delineator and the result of the algorithm should be discarded in preference to the
manual result.
This chapter contains four principal conclusions that have been drawn from the work carried out. The original aim of this work is to improve the precision of wound measurement by removing the subjective elements of human observation and manual dexterity. The properties of most wound images prevent many deterministic image processing methods from offering a viable alternative to the established manual methods, and instead attention has been focussed on active contour models to provide a mechanism for refining manual wound area measurements.

8.1 Manual Delineation Performance

The manual wound measurement results of Chapter 4 show that not all of the delineators involved in the trial measure the same wound area. This is due to differences of opinion concerning which parts of an image belong to a wound. Also, significant differences in measurement precision show that some delineators generally achieve a higher level of repeatability in their measurements. The precision level varies with the delineator, the wound image itself, and the size of the display of the image on a monitor. Images of wounds that appear well defined have relatively strong edge strength at the wound boundary. The last variable, the size of image displayed, means that the precision of measurements made on smaller wounds improves with magnification.
8.2 Algorithm Performance

Active contour models are appropriate for the wound measurement problem because they require an initial estimate of the final boundary, which can be provided by a trained delineator. Thus the solution is obtained by local modification of the contour, without regard to general properties of the enclosed region or of the surrounding tissue. Four different formulations of active contour model algorithms have been tested. The TN algorithm has proved to be the most appropriate single algorithm for improving wound measurements:

- It significantly reduces variance in 90% of test cases.
- It produces the highest number of agreements with manual measurements (60% of test cases).

The GA and MG algorithms, whilst generally inferior in performance to the TN algorithm supersede it when the image gradient information becomes too ambiguous or when the initial filter scale is too high. In such cases the gray level term of the MG algorithm replaces the lost information, and the GA algorithm benefits from retaining the initial contour in the iteration equations. The setting of contour regularisation parameters to apply across a wide range of images causes precision loss in the GA and TN algorithms. This requires a method for adapting the parameters to each image without compromising the stability of the solution. The parameterless MX algorithm, which is adaptive in this sense, does not perform as well as the TN algorithm.

Automation of the algorithms has not been achieved because it is considered that not enough reliable information generally exists in wound images in general to allow automatically initialised contours or region growing contours to be deformed to a reliable solution.
8.3 General Statement

This thesis claims the following innovations as part of its original contribution to knowledge:

- The extension of the minimax principle, as applied to active contour models, to the continuous spatial domain and its implementation in the form of second-order continuous ($C^2$) cubic B-Spline curve elements.
- The definition in the continuous domain of a scale-invariant curvature term (equation 3.22) and its implementation with $C^2$ cubic B-Splines.
- The definition of a gray level external energy term that allows an active contour model to adaptively seek a nearby contour of greatest homogeneity with regard to the variations of gray level along that contour (equation 3.26).
- The modification of the variational forces arising from the external energy by removing the force component that acts tangentially to the contour (implemented in the TN algorithm — cf. §3.3.1). The result is an external force that is oriented normal to the contour, and this feature removes the tendency for the elements to shrink at points of high gradient.
- The formulation in the continuous spatial domain of an elastic force that acts only tangentially to the contour and its implementation in a finite element active contour model using an approximation step (cf. 3.3.3).
- The development and testing of (a) an ANOVA model for determining the existence of bias between area measurements from different delineators, and (b) linear regression models for estimating delineator performance as a function of display size.
- The development of a graphical method for the simultaneous display of the significance of unequal variance $t$-tests and variance ratio tests and its use to determine the performance of the four active contour model algorithms with respect to the performance criteria stipulated in §6.1.
The algorithms described in this work, their implementation and testing has resulted in a useful wound measurement tool. The work on parameter setting in Chapter 5 enables the algorithms to operate with little end-user intervention. Thus they may be used in a clinical setting by a physician and do not require technical or expert intervention. The algorithms produce improved precision area measurements, the only human effort required being an initial, approximate delineation of the wound boundary using a mouse.

The image and delineator-conditional ability of the algorithms to improve precision should be weighed carefully on a case-by-case basis, so that when complex or poorly defined wounds are measured or where poorly posed and poorly illuminated video capture is apparent the expectation for improved precision may not be fulfilled.

Since so many cases (40%) produce bias errors, a plan for wound measurement that accounts for this factor is therefore recommended:

- The delineator makes several delineations upon the same image in order to improve the standard error of the area measurement.

- All of the previous images, delineations and area measurements are stored and at least the most recent ones recalled at the next measurement/treatment time for the purposes of comparison. One reason for keeping a visual record is that sometimes two wounds may join to produce one large wound and ignorance of this fact could lead to large increases in size being recorded.

- The gradient image of the wound is presented to the delineator for inspection. If this does not contain a recognisable path for the edge of the wound the delineator can opt to use another algorithm or exclude its use altogether. This does however, introduce a subjective dimension.
8.4 Limitations of this Work

There are many possible sources of error in video measurements (Plassmann 1992). In terms of experimental analysis the most obvious missing component of the work is an investigation into the errors arising from taking measurements from different images of the same wound taken at the same visit. The measurement of image area has concentrated on the 2-D projection of the object onto the camera plane. Further work is necessary therefore to estimate the sum contribution of these error sources and weigh their effect upon the reliability of measurements.

There is much scope for further investigation in the area of contour initialisation. Chapter 2 discussed the various methods reported in practice and introduced some ideas. However, the algorithms reported in this thesis have not employed such methods due to the requirement that they work across a broad range of wound images. This could well involve a large amount of work determining amongst other things, the region, both inside and outside of the wound boundary, within which a contour can be placed and successfully converge to a solution with an error on a par with the results quoted in Chapter 6.

8.5 Recommendations for Further Work

Since bias errors between manually defined areas and algorithm defined areas are the main source of error it is recommended that a further study be performed which tracks the progress of a set of wounds. The same procedures for measurement as described in this work may then be used to determine whether or not the bias errors can be wholly or partly ignored, or whether they introduce a further degree of variability. If the latter case is found to be true, then a systematic error in one set of measurements will become essentially a random error when considered over many such sets of measurements, i.e. the bias between manual and algorithmic measurements will be unpredictable from one wound inspection to the next.
A further idea would be to employ the multiple contour principle of Gunn and Nixon (1994) so that several starting estimates would be supplied for a single measurement. Instead of these contours being mutually independent, the multiple contour formulation would enforce a degree of similarity between results, thereby reducing variability and thus increasing precision. Furthermore, it may be possible, where more than two contours are used, to consider energies for contour segments that are in closer agreement and to exclude any significantly different contour segments. The multiple contour configuration should not require so much dependence upon scale descent techniques to improve precision and thus should not suffer so much from the effect of bias.
FORMULATION OF B-SPLINE AND HERMITE FINITE ELEMENT MODELS

The formulation of the finite element active contour models described in Chapters 2 and 3 relies upon the specification of shape functions defining the piecewise continuous curve. The cubic B-Spline shape functions and the cubic Hermite shape functions are defined in A.1. The resulting stiffness matrix for each set of shape functions is derived in A.2.

A.1 Specification of Finite Element Shape Functions

Cubic Hermite Shape Functions

The set of four cubic hermite blending functions are defined as (Bartels et al., 1987):

\[
\begin{align*}
  h_{00}(s) &= 2s^3 - 3s^2 + 1 \\
  h_{10}(s) &= -2s^3 + 3s^2 \\
  h_{01}(s) &= s^3 - 2s^2 + s \\
  h_{11}(s) &= s^3 - s^2 
\end{align*}
\]

(1)

Define a vector of the above shape functions:

\[
h(s) = [h_{00}(s) \ h_{10}(s) \ h_{01}(s) \ h_{11}(s)]
\]

(2)

The vector of control vertices defining one element of a piecewise parametric curve required for use with these equations is given by:

\[
V_e = [V_n \ V_{n+1} \ D_n \ D_{n+1}]^T
\]

(3)
where the end-derivatives, \( D_n \), are specified as the default slopes:

\[
D_n = \frac{1}{2} (V_{n+1} - V_{n-1})
\]

Thus any point in one element of the curve is defined by:

\[
v_e(s) = h(s)V_e
\]

It is helpful at this stage to rewrite (5) making use of the definition for \( D_n \) and \( D_{n+1} \) given by (4). Thus the matrix equation for a point on the curve may be written explicitly in terms of the control vertices. This provides a modified set of cubic Hermite shape functions:

\[
H(s) = \begin{bmatrix}
    H_0(s) \\
    H_1(s) \\
    H_2(s) \\
    H_3(s)
\end{bmatrix}
= \begin{bmatrix}
    -\frac{1}{2}h_0(s) \\
    h_{00}(s) - \frac{1}{2}h_{11}(s) \\
    h_{01}(s) + \frac{1}{2}h_{10}(s) \\
    \frac{1}{2}h_{11}(s)
\end{bmatrix}
= \begin{bmatrix}
    -\frac{1}{2}s^3 + s^2 - \frac{1}{2}s \\
    \frac{1}{2}s^3 - \frac{1}{2}s^2 + 1 \\
    -\frac{1}{2}s^3 + 2s^2 + \frac{1}{2}s \\
    \frac{1}{2}s^3 - \frac{1}{2}s^2
\end{bmatrix}
\]

To be clear, the corresponding vector of control vertices is given by:

\[
V_e = [V_{n-1} \ V_n \ V_{n+1} \ V_{n+2}]^T
\]

Thus the curve defined in (5) is now equivalently specified by the inner product of (6) and (7). This results in a more manageable form that will aid in the development of the stiffness matrix arising from the definition of the internal energy terms of the active contour model.
Cubic B-Spline Shape Functions

The cubic B-Spline shape functions are defined by (Bartels et al., 1987):

\[
B(s) = \begin{bmatrix}
  b_{-3}(s) \\
  b_{-2}(s) \\
  b_{-1}(s) \\
  b_{-0}(s)
\end{bmatrix} = \frac{1}{6} \begin{bmatrix}
  1 - 3s + 3s^2 - s^3 \\
  4 - 6s^2 + 3s^3 \\
  1 + 3s + 3s^2 - 3s^3 \\
  s^3
\end{bmatrix}
\]

Thus any point on the curve \(v(s)\) contained within an element \(e\) is given by:

\[
v_e(s) = B(s)V_e
\]

Note that the cubic curve defined by set of control vertices using B-Spline elements will differ from the cubic curve defined by the same set of control vertices when the elements are hermites. Hermite interpolation allows the curve to pass through the control vertices which are located at the end points of each element. B-Splines do not possess this ability to interpolate the control vertices so that in general the control vertices and element end points do not coincide.

A.2 Development of Finite Element Matrix Equations

Application of the Rayleigh-Ritz finite element method with cyclic boundary conditions and constant internal parameters:

(1) With reference to the energy functional (2.1) in Chapter 2: for each element, \(e\), derive the following

\[
\frac{\partial}{\partial X_e} \left( E^*_e \right) = \int_{s=0}^{s=1} \frac{\partial}{\partial X_e} \left( \frac{a}{2} (N_s(s)X_e)^2 + \frac{b}{2} (N_{ss}(s)X_e)^2 \right) ds + \int_{s=0}^{s=1} \frac{\partial}{\partial X_e} \left( P(V_e) \right) ds
\]

(10.1)
where $X_e$ and $Y_e$ are the first and second columns of the elemental control-vertex vector, $V_e$, respectively.

and $N_s(s)$ and $N_{ss}(s)$ respectively are matrices of the first and second derivatives of the elemental shape functions (regardless of specific form).

(2) Performing the differentiation with respect to the control vertices in (10.1) and (10.2) yields:

\[
\frac{\partial E^*_e}{\partial X_e} = \int_0^1 \alpha (N_s^T N_s) X_e + \beta (N_{ss}^T N_{ss}) X_e ds + \int_0^1 \frac{\partial}{\partial X_e} P(V_e) ds
\]

\[
\frac{\partial E^*_e}{\partial Y_e} = \int_0^1 \alpha (N_s^T N_s) Y_e + \beta (N_{ss}^T N_{ss}) Y_e ds + \int_0^1 \frac{\partial}{\partial Y_e} P(V_e) ds
\]

(11.1) and (11.2) may be combined into a single matrix equation:

\[
\frac{dE^*_e}{dV_e} = \int_0^1 \left( \alpha (N_s^T N_s) + \beta (N_{ss}^T N_{ss}) \right) V_e ds + \int_0^1 \nabla P(V_e) ds
\]

Without regard to the specific form of shape equations to be used, (12) will reduce to the form of the following equation:

\[
\frac{dE^*_e}{dV_e} = (\alpha K_1 + \beta K_2) V_e + f(V_e) = K_e V_e + f_e
\]

Clearly, to proceed any further with (12) it is necessary to specify the form of the shape functions. This will allow direct evaluation of the first integral term on the RHS of (12).
Due to the complex nature of the external potential function, $P$, the integration of its partial spatial derivatives along each element must be performed numerically. Thus no further reduction of $f_e$ is possible.

### A.2.1 Hermite Cubic Elemental Stiffness Matrix

Substituting the cubic hermite shape functions from (6) for $N(s)$ in (12) and performing the necessary calculus operations yields a matrix equation in the form of (13) where $K_1$ and $K_2$ have the following specific definitions:

\[
K_1 = \frac{1}{120} \begin{bmatrix}
4 & -7 & 2 & 1 \\
-7 & 136 & -131 & 2 \\
2 & -131 & 136 & -7 \\
1 & 2 & -7 & 4
\end{bmatrix}
\]

\[
K_2 = \frac{1}{120} \begin{bmatrix}
120 & -300 & 240 & -60 \\
-300 & 840 & -780 & 240 \\
240 & -780 & 840 & -300 \\
-60 & 240 & -300 & 120
\end{bmatrix} = \frac{1}{2} \begin{bmatrix}
2 & -5 & 4 & -1 \\
-5 & 14 & -13 & 4 \\
4 & -13 & 14 & -5 \\
-1 & 4 & -5 & 2
\end{bmatrix}
\]

Using the definition $K_e = \alpha K_1 + \beta K_2$ for each element $e$ in the domain of the solution, one may deduce that the final system assemblage stiffness matrix will be a cyclic sparse matrix (hepta-diagonal), the rows of which will be formed by shifted versions of the following row vector:

\[
K_{row} = \frac{\alpha}{120} \left[ \cdots 0 \ 1 \ 4 \ -145 \ 280 \ -145 \ 4 \ 1 \ 0 \ \cdots \right] + \frac{\beta}{120} \left[ \cdots 0 \ -60 \ 480 \ -1380 \ 1920 \ -1380 \ 480 \ -60 \ 0 \ \cdots \right]
\]
A.2.2 Cubic B-Spline Shape Functions

In an identical manner to the above, substituting the cubic B-Spline shape functions (8) into (12) yields:

\[
K_1 = \frac{1}{120} \begin{bmatrix}
6 & 7 & -12 & -1 \\
7 & 34 & -29 & -12 \\
-12 & -29 & 34 & 7 \\
-1 & -12 & 7 & 1
\end{bmatrix}
\]

\[
K_2 = \frac{1}{120} \begin{bmatrix}
40 & -60 & 0 & 20 \\
-60 & 120 & -60 & 0 \\
0 & -60 & 120 & -60 \\
20 & 0 & -60 & 40
\end{bmatrix}
= \frac{1}{6} \begin{bmatrix}
2 & -3 & 0 & 1 \\
-3 & 6 & -3 & 0 \\
6 & 0 & -3 & 6 \\
1 & 0 & -3 & 2
\end{bmatrix}
\]

As in the previous section, the B-Spline equivalent of (16) becomes:

\[
K_{row} = \frac{\alpha}{120} \left[ \cdots 0 -1 -24 -15 80 -15 -24 -1 0 \cdots \right] + \frac{\beta}{120} \left[ \cdots 0 20 0 -180 320 -180 0 20 0 \cdots \right]
\]
## MANUAL DELINEATION

### EXPERIMENTAL DATA

<table>
<thead>
<tr>
<th>Image</th>
<th>Delineator mean areas $A_i$ [pixels]</th>
<th>Grand mean $A$ [pixels]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>8372</td>
<td>8229</td>
</tr>
<tr>
<td>2</td>
<td>6256</td>
<td>6395</td>
</tr>
<tr>
<td>3</td>
<td>45640</td>
<td>46550</td>
</tr>
<tr>
<td>4</td>
<td>34470</td>
<td>35900</td>
</tr>
<tr>
<td>5</td>
<td>29370</td>
<td>30160</td>
</tr>
<tr>
<td>6</td>
<td>23750</td>
<td>25250</td>
</tr>
<tr>
<td>7</td>
<td>3968</td>
<td>4024</td>
</tr>
<tr>
<td>8</td>
<td>35050</td>
<td>37720</td>
</tr>
<tr>
<td>9</td>
<td>44240</td>
<td>45760</td>
</tr>
<tr>
<td>10</td>
<td>43700</td>
<td>47000</td>
</tr>
</tbody>
</table>

*Table B.1* Mean areas for each delineator and 'grand mean' for ANOVA study in §4.1.

<table>
<thead>
<tr>
<th>Image</th>
<th>Delineator area variances $s^2_i$ [pixels$^2$]</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>1</td>
<td>138800</td>
<td>88860</td>
</tr>
<tr>
<td>2</td>
<td>24690</td>
<td>15410</td>
</tr>
<tr>
<td>3</td>
<td>1153200</td>
<td>439200</td>
</tr>
<tr>
<td>4</td>
<td>842600</td>
<td>254500</td>
</tr>
<tr>
<td>5</td>
<td>8524000</td>
<td>470400</td>
</tr>
<tr>
<td>6</td>
<td>608800</td>
<td>237700</td>
</tr>
<tr>
<td>7</td>
<td>38060</td>
<td>16570</td>
</tr>
<tr>
<td>8</td>
<td>7998000</td>
<td>585900</td>
</tr>
<tr>
<td>9</td>
<td>3528000</td>
<td>1629000</td>
</tr>
<tr>
<td>10</td>
<td>669900</td>
<td>1907000</td>
</tr>
</tbody>
</table>

*Table B.2* Area variance estimates for each delineator and pooled estimate (MSE) for ANOVA study and equality of variance test in §4.1. Values of MSE in parentheses indicate unequal delineator variances.
ACTIVE CONTOUR MODEL
EXPERIMENTAL DATA

<table>
<thead>
<tr>
<th>Image</th>
<th>Biases for Algorithms $B_w$ [% of Manual Area]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>GA</td>
</tr>
<tr>
<td>1.1</td>
<td>-3.9</td>
</tr>
<tr>
<td>1.2</td>
<td>0.1</td>
</tr>
<tr>
<td>1.3</td>
<td>3.4</td>
</tr>
<tr>
<td>1.4</td>
<td>3.0</td>
</tr>
<tr>
<td>1.5</td>
<td>1.8</td>
</tr>
<tr>
<td>1.6</td>
<td>1.9</td>
</tr>
<tr>
<td>1.7</td>
<td>-3.2</td>
</tr>
<tr>
<td>1.8</td>
<td>2.0</td>
</tr>
<tr>
<td>1.9</td>
<td>-4.3</td>
</tr>
<tr>
<td>1.10</td>
<td>6.6</td>
</tr>
<tr>
<td>2.1</td>
<td>-1.6</td>
</tr>
<tr>
<td>2.2</td>
<td>-4.1</td>
</tr>
<tr>
<td>2.3</td>
<td>0.8</td>
</tr>
<tr>
<td>2.4</td>
<td>4.7</td>
</tr>
<tr>
<td>2.5</td>
<td>-1.1</td>
</tr>
<tr>
<td>2.6</td>
<td>-0.4</td>
</tr>
<tr>
<td>2.7</td>
<td>1.5</td>
</tr>
<tr>
<td>2.8</td>
<td>4.8</td>
</tr>
<tr>
<td>2.9</td>
<td>-3.8</td>
</tr>
<tr>
<td>2.10</td>
<td>-14.5</td>
</tr>
</tbody>
</table>

Table C.1 Biases between for manual and algorithm area measurement samples (cf. §6.4). In this study, the mean of manually delineated area measurements for each wound (see Appendix B) is used as the reference. The values were calculated using equation (6.11).
Table C.2 Area measurement precision values for algorithm area measurement samples (cf. §6.4).

In this study, the mean of manually delineated area measurements for each wound (see Appendix B) is used as the reference. The values were calculated using equation (6.11).
REFERENCES

Aarts, E. H. L. and van-Laarhoven, P. J. M.
Simulated Annealing: A Pedestrian Review of the Theory and Some Applications
in Pattern Recognition Theory and Applications (Eds.: Devijver, P. A. and Kittler, J.)
NATO Advanced Science Institutes Series, Springer-Verlag, 1987, pp. 179-192

Adams, R. and Bischof, L.
Seeded Region Growing
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 16, No. 6, June 1994, pp. 641-647

Amini, A. A., Weymouth, T. E. and Jain, R. C.
Using Dynamic Programming for Solving Variational Problems in Vision
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 12, No. 9, September 1990, pp. 855-867

Anderson, V. L. and McLean, R. A.
Design of Experiments: A Realistic Approach
Marcel Dekker, 1974

Ballard, D. H. and Brown, C. M.
Computer Vision
Prentice-Hall, 1982

Barford, N. C.
Experimental Measurements: Precision, Error and Truth

Bartels, R. H., Beatty, J. C. and Barsky, B. A.
An Introduction to Splines for use in Computer Graphics and Geometric Models
Morgan Kaufmann, 1987

Benson, A. and Evans, D. J.
A Normalized Algorithm for the Solution of Positive Definite Symmetric Quindiagonal Systems of Linear Equations
ACM Trans. Math. Software
Vol. 3, March 1977, pp. 96-103

Bergholm, F.
Edge Focusing
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 9, No. 6, June 1987, pp. 726-741

Berris, W. P. and Sangwine, S. J.
A Colour Histogram Clustering Technique for Tissue Analysis of Healing Skin Wounds
Proc. 5th International Conference on Image Processing and its Applications
1997, pp. 693-697
Box, G. E. P.
Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance Problems – I
Vol. 25, 1954, pp. 290-302

Canny, J.
A Computational Approach to Edge Detection
*IEEE Trans. Pattern Analysis and Machine Intelligence*
Vol. 8, No. 6, June 1986, pp. 679-698

Chen, P. C. and Pavlidis, T.
Segmentation by Texture using a Co-Occurrence Matrix and A Split-and-Merge Algorithm
*Computer Vision, Graphics and Image processing*
Vol. 10, 1979, pp. 172-182

Chou, J.-S., Chen, C.-T., Chen, S.-Y. and Lin, W.-C.
Parallel Implementation of an Adaptive Split-and-Merge Method for Medical Image Segmentation
*SPIE Medical Imaging VI: Image Processing*

Chow, C. K. and Kaneko, T.
Automatic Boundary Detection of the Left Ventricle from Cineangiograms
*Computers and Biomedical Research*
Vol. 5, 1972, pp. 388-410

Cohen, L. D.
On Active Contour Models and Balloons
*Computer Vision Graphics and Image Processing: Image Understanding*

Cohen, L. D. and Cohen, I.
Finite-Element Methods for Active Contour Models and Balloons for 2-D and 3-D Images
*IEEE Trans. Pattern Analysis and Machine Intelligence*
Vol. 15, No. 11, November 1993, pp. 1131-1147

Däschlein, R., Waschulzik, T., Brauer, W., Arnoldi, H.-M. and Englmeier, K.-H.
Segmentation of Computertomographies with Neural Networks based on Local Features
*Proc. Int. Conf. on Neural Networks and Expert Systems in Medicine and Healthcare*
NNESMED 94, Plymouth, UK, 1994, pp. 240-247

Davis, M. H., Dunkley, P., Harden, R. M., Harding, K., Laidlaw, J. M., Morris, A. M., and
Wood, R. A. B.
*The Wound Programme*
Centre for Medical Education, University of Dundee, 1992, p. 123

Deming, W. E.
*Some Theory of Sampling*
Wiley, 1950

Denton, W. E., Duller, A. W. G. and Fish, P. J.
Boundary Detection for Skin Lesions: An Edge Focusing Algorithm
*IEE Conf. Image Processing and its Applications*
REFERENCES

Deriche, R.
Using Canny’s Criteria to Derive a Recursively Implemented Optimal Edge Detector
*International Journal of Computer Vision*

Duda, R. O. and Hart, P. E.
*Pattern Classification and Scene Analysis*
Wiley, 1973

Farlow, S. J.
*Partial Differential Equations: for Scientists and Engineers*
Wiley, 1982

Fu, K. S. and Mui, J. K.
A Survey on Image Segmentation
*Pattern Recognition*
Vol. 13, 1981, pp. 3-10

Gauch, J. M.
Image Segmentation and Analysis via Multiscale Gradient Watershed Hierarchies
*IEEE Trans. Image Processing*
Vol. 8, No. 1, January 1999, pp. 69-79

Gennert, M. A. and Yuille, A. L.
Determining the Optimal Weights in Multiple Objective Function Optimization
*Proc. 2nd International Conference on Computer Vision*
1988, pp. 87-89

Golston, J. E., Moss, R. H. and Stoecker, W. V.
Boundary Detection in Skin Tumour Images: an Overall Approach and a Radial Search Algorithm
*Pattern Recognition*
Vol. 23, No. 11, 1990, pp. 1235-1247

Gonzalez, R. C. and Woods, R. E.
*Digital Image Processing*
Addison-Wesley, 1992

Gunn, S. R. and Nixon, M. S.
A Dual Active Contour
*Proc. 5th British Machine Vision Conference*
September 1994, York, UK, pp. 305-314

Haralick, R. M.
Zero Crossing of Second Directional Derivative Edge Operator
*SPIE Symposium on Robot Vision* (Washington DC)

Haralick, R. M. and Shapiro, L. G.
Survey: Image Segmentation Techniques
*Computer Vision, Graphics and Image Processing*
Vol. 29, 1985, pp. 100-132
Haris, K., Efstratiadis, S. N., Maglaveras, N. and Katsaggelos, A. K.
Hybrid Image Segmentation Using Watersheds and Fast Region Merging
*IEEE Trans. Image Processing*
Vol. 7, No. 12, December 1998, pp. 1684-1699

Healey, G.
Segmenting Images Using Normalized Color
*IEEE Trans. Systems, Man, and Cybernetics*
Vol. 22, No. 1, January/February 1992, pp. 64-73

Horowitz, S. L. and Pavlidis, T.
Picture Segmentation by a Tree Traversal Algorithm
*Journal of the Association of Computing Machinery*
Vol. 23, 1976, pp. 368-388

Ivins, J. and Porril, J.
Statistical Snakes: Active Region Models
*Proc. 5th British Machine Vision Conference*
1994, pp. 377-386

Jackson, T. R., Brookeman, J. R. and Merickel, M. B.
Classification of Medical Images using Context Dependent Methods
*Proc. SPIE (Image Processing)*
Vol. 1898, 1993, pp. 430-441

James, M.
*Pattern Recognition*
Oxford: BSP Professional, 1987

Jarkrans, T., Bengtsson, E., Eriksson, O., Nordin, B. and Stenkvist, B.
Using Tracking Algorithms for Cell Segmentation
*Proc. 1st Scandinavian Conference on Image Analysis*
January 1980, pp. 325-330

Johnson, R. A. and Bhattacharyya, G. K.
*Statistics: Principles and Methods*

Jones, B. F. and Plassmann, P.
An Instrument to Measure the Dimensions of Skin Wounds
*IEEE Trans. Biomedical Engineering*

Karaolani, P., Sullivan, G. D. and Baker, K. D.
Parabolic and Hermite Cubic Finite Elements: A Flexible Technique for Deformable Models
*Proc. 1st British Machine Vision Conference*
1990, pp. 325-330

Karaolani, P., Sullivan, G. D. and Baker, K. D.
Active Contours using Finite Elements to Control Local Scale
*Proc. 3rd British Machine Vision Conference*
1992, pp. 481-487
Kass, M., Witkin, A. and Terzopoulos, D.
Snakes: Active Contour Models
Proc. 1st International Conference on Computer Vision
1987, pp. 259-268

Kirk, R. E.
Experimental Design: Procedures for the Behavioural Sciences
Brooks-Cole, 1968

Kittler, J., Illingworth, J., Foglein, J. and Paler, K.
An Automatic Thresholding Algorithm and its Performance
IEEE Proc. 7th International Conference on Pattern Recognition
Vol. 1, 1984, pp. 287-289

Kittler, J. and Illingworth, J.
On Threshold Selection Using Clustering Criteria
IEEE Trans. Systems, Man, and Cybernetics
Vol. 15, No. 5, September/October 1985, pp. 652-655

Kittler, J. and Illingworth, J.
Minimum Error Thresholding
Pattern Recognition
Vol. 19, No. 1, 1986, pp. 41-47

Klinger, A.
Data Structures and Pattern Recognition
Proc. 1st International Joint Conference on Pattern Recognition
Washington, D. C., October 1973, pp. 497-498

Kohler, R.
A Segmentation System Based on Thresholding
Computer Graphics and Image Processing
Vol. 15, 1981, pp. 319-338

Kundin, J. I.
A New Way to Size up a Wound
Amer. J. Nursing

Lai, K. F. and Chin, R. T.
On Regularisation, Formulation and Initialisation of the Active Contour Models (Snakes)
Asian Conference on Computer Vision
1993, pp. 542-545

Lai, K. F. and Chin, R. T.
Deformable Contours: Modeling and Extraction
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 17, 1995, pp. 1084-1090

Laws, K. I.
Texture Energy Measures
DARPA Image Understanding Workshop (Los Angeles)
1979, pp. 47-51
References

Two Graph Searching Techniques for Boundary Finding in White Blood Cell Images
Vol. 8, 1978, pp. 293-308

Leymarie, F. and Levine, M. D.
Shape Features Using Curvature Morphology
Vol. 1192, Part 2, Nov. 1989, pp. 536-547

Leymarie, F. and Levine, M. D.
Tracking Deformable Objects in the Plane Using an Active Contour Model
*IEEE Trans. Pattern Analysis and Machine Intelligence*
Vol. 15, No. 6, June 1993, pp. 617-634

Marr, D. and Hildreth, E.
Theory of Edge Detection
*Proc. Royal Society*

Martelli, A.
Edge Detection Using Heuristic Search Methods
*Computer Graphics and Image Processing*
Vol. 1, 1972, pp. 169-182

Montgomery, D. C.
*Design and Analysis of Experiments*

Muerle, J. and Allen, D.
Experimental Evaluation of Techniques for Automatic Segmentation of Objects in a Complex Scene
in *Pictorial Pattern Recognition* (Eds.: Cheng, G., et al.)
Thompson, 1968, pp. 3-13

Neave, H. R. and Worthington, P. L.
*Distribution-Free Tests*
Unwin-Hyman, 1988

Neter, J., Kutner, M. H., Nachtsheim, C. J. and Wasserman, W.
*Applied Linear Statistical Models*

Niblack, W.
*An Introduction to Digital Image Processing*
Prentice-Hall, 1986

Nilsson, N. J.
*Problem-Solving Methods in Artificial Intelligence*
McGraw-Hill, 1971
References

O'Gorman, L., Sanderson, A. C., Preston, K. Jr. and Dekker, A.
Image Segmentation and Nucleus Classification for Automated Tissue Section Analysis
Proc. IEEE Computer Vision Pattern Recognition
1983, pp. 89-94

Ohlander, R., Price, K. and Reddy, D. R.
Picture Segmentation Using A Recursive Region Splitting Method
Computer Graphics and Image Processing
Vol. 8, 1978, pp. 313-333

Ohta, Y-U., Kanade, T. and Sakai, T.
Color Information for Region Segmentation
Computer Graphics and Image Processing
Vol. 13, 1980, pp. 222-241

Ossen, A., Zamzow, T., Oswald, H. and Fleck, E.
Segmentation of Medical Images using Neural-Network Classifiers
Proc. Int. Conf. on Neural Networks and Expert Systems in Medicine and Healthcare
NNESMED 94, Plymouth, UK, 1994, pp. 427-432

Otsu, N.
A Threshold Selection Method from Gray Level Histograms
IEEE Trans. Systems, Man, and Cybernetics
Vol. 9, 1979, pp. 62-66

Perona, P. and Malik, J.
Scale-Space and Edge Detection Using Anisotropic Diffusion
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 12, No. 7, July 1990, pp. 629-639

Plassmann, P.
Measuring Area and Volume of Leg Ulcers by Structured Light

Plassmann, P.
Measuring Wounds – A Guide to the Use of Wound Measurement Techniques
J. Wound Care
Vol. 4, No. 6, June 1995, pp. 269-272

Plassmann, P. and Jones, T. D.
MAVIS – A Non-Invasive Instrument to Measure Area and Volume of Wounds
Medical Engineering and Physics
Vol. 20, 1998, pp. 325-331

Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P.
Numerical Recipes in C

Radeva, P. and Marti, E.
An Improved Model of Snakes for Model-Based Segmentation
Proc. 6th Int. Conf. Computer Analysis of Images and Patterns (CAIP ‘95, Prague)
Springer-Verlag, September 1995, pp. 515-520
References

Ranganath, S.
Analysis of the Effects of Snake Parameters on Contour Extraction
Proc. 2nd Int. Conf. Automation, Robotics and Computer Vision
Singapore, September 1992, pp. CV-4.5.1 – CV-4.5.5

Ruiz, E. E. S. and Fairhurst, M. C.
Two-Dimensional Echocardiographic Images
Vol. 142, No. 3, June 1995, pp. 121-127

Sahoo, P. K., Soltani, S., Wong, A. K. C. and Chen, Y. C.
A Survey of Thresholding Techniques
Computer Vision, Graphics and Image Processing
Vol. 41, 1988, pp. 233-260

Schmid, P. and Fischer, S.
Colour Segmentation for the Analysis of Pigmented Skin Lesions
Proc. 5th International Conference on Image Processing and its Applications
July 1997, pp. 688-692

Staib, L. H. and Duncan, J. S.
Boundary Finding with Parametrically Deformable Models
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 14, No. 11, November 1992, pp. 1061-1075

Stasa, F. L.
Applied Finite Element Analysis for Engineers
CBS, 1985

Storvik, G.
A Bayesian Approach to Dynamic Contours through Stochastic Sampling and Simulated Annealing
IEEE Trans. Pattern Analysis and Machine Intelligence
Vol. 16, No. 10, October 1994, pp. 976-986

Stroud, K. A.
Engineering Mathematics

Sullivan, G. D., Worrall, A. D., Hockney, R. W. and Baker, K. D.
Active Contours in Medical Image Processing using a Networked SIMD Array Processor
Proc. 1st British Machine Vision Conference
1990, pp. 395-400

Terzopoulos, D.
On Matching Deformable Models to Images
in Topical Meeting Machine Vision Technological Digest Series (Washington DC)

Terzopoulos, D.
Matching Deformable Models to Images: Direct and Iterative Solutions
in Topical Meeting Machine Vision Technological Digest Series (Washington DC)
Optical Soci. Amer., Vol. 12, March 1987, pp. 164-167
Therrien, C. W.
*Decision, Estimation and Classification*
Wiley, 1989

Topping, J.
*Errors of Observation and their Treatment*

Umbaugh, S. E., Moss, R. H., Stoecker, W. V. and Hance, G. A.
Automatic Color Segmentation Algorithms
*IEEE Engineering in Medicine and Biology*
September, 1993, pp. 75-82

Wadsworth, H. M.
*Handbook of Statistical Methods for Engineers and Scientists*

Wahl, F. M.
*Digital Image Processing*
Artech House, 1987

Weszka, J.
A Survey of Threshold Selection Techniques
*Computer Graphics and Image Processing*

Williams, D. J. and Shah, M.
A Fast Algorithm for Active Contours and Curvature Estimation
*CVGIP: Image Understanding*

Wu, K., Gauthier, D. and Levine, M. D.
Live Cell Image Segmentation
*IEEE Trans. Biomedical Engineering*
Vol. 42, No. 1, January 1995, pp. 1-12

Xu, G., Segawa, E. and Tsuji, S.
Robust Active Contours With Insensitive Parameters
*Pattern Recognition*
Vol. 27, No. 7, 1994, pp. 879-884