COMPUTER AIDED DETECTION OF NATURAL REFERENCE MARKERS IN SERIAL RADIOGRAPHS OF GROWING BONE

STEPHEN DAVID SPRINGATE
B.D.S.Hons., M.Sc., F.D.S., M.Orth.R.C.S.(Eng)
Department of Orthodontics
Institute of Dental Surgery
University of London

Thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
Faculty of Medicine
University of London
1993

ABSTRACT

This thesis describes the theoretical and experimental development and testing of a computer-based image processing system for the detection of stable structural features in serial radiographs of growing bone.

The study is divided into three parts. First, a simple theoretical model for the detection of stable structures in a radiographic sequence was developed together with a more detailed statistical model of the processes which hinder detection. Analysis of the models indicated that the processing procedures needed to perform the detection were those described by the so-called "matched filter equation".

Secondly, the assumptions of the model were tested in a series of experiments using serial radiographs of 24 children with implanted tantalum markers, drawn from the files of the University of Washington, USA. Information gained from these experiments prompted changes to the detection method to allow the radiographic data to comply with these assumptions and provide robustness against image noise.

The method was implemented as a semi-automatic image processing system on a Sun 3/160 computer. The system was tested in a further series of experiments using a second sample of radiographs of 28 children with implanted markers and with radiographs of dried human skulls.

The results of these experiments indicated that the system could detect persistent structural features whose positions were consistent with the stable tantalum markers and that the system was robust against minor changes in projection between radiographs.

Finally, the detection system was applied to the investigation of mandibular incisor eruption from intra-alveolar to pubertal eruption stages using serial lateral cephalometric radiographs of 11 children. This investigation revealed two new features of incisor eruption: first, the eruption paths deviated lingually at, or just prior to, alveolar emergence; and second, post-emergence eruption was characterised by a labio-lingual oscillation of the crowns of the incisors.

ACKNOWLEDGMENTS

The work which led to the writing of this thesis could never have taken place without the help, encouragement and support of many people, too many to properly acknowledge them all.

I particularly wish to express my gratitude to the three people who made this study possible: the late Dr J. Rodney Mathews and Professor Sheldon Baumrind of the Department of Growth and Development, University of California at San Francisco and Professor Robert Little of the University of Washington, whose generosity I can never hope to repay.

I am forever indebted to Professor Michael Duff and his staff whose kindness and understanding of the problem finally set this work in motion. I especially wish to thank to Dr Terry Fountain and Dr Mike Forshaw, for the many hours they spent listening and encouraging, and Dr David Greenwood, for writing the original correlation program (and for helping me believe that I knew some physics). I also warmly thank Dr Charles Wiles of the Planetary Sciences Group for allowing me to adopt and adapt his correlation detection program.

I am grateful to Dr Denis Rosen for his help and guidance in defining what I needed to do. I am indebted to Dr Yushido Bando for Japanese translations, to Elsa Ekström, for Danish and German translation, to Rosemary Bowker for assistance in recording radiographs of dried bone and to Dr John Bulman for statistical advice. I would also like to thank my good friend Steve Jones for his meticulous proof reading.

I would like to express my sincere thanks to my supervisors, Professor Murray Meikle and the late Professor William Houston whose death robbed me not only of a supervisor but also a valued friend. I am deeply grateful to Dr Nigel Hunt who stepped in to bridge the chasm left by the loss of Professor Houston and whose help and encouragement are greatly appreciated.

Finally, I thank my family to whom I have given so little of my time during the preparation of this work - thank you for indulging me.

CONTENTS

CORRIGENDA

p.28	line 29	... proportional ...	should read	... increases with ...
p.29	line 26	... proportional ...	should read	... inversely related ...
p.60	line 1	... most ...	should read	... many ...
p.65	line 9	... impractical ...	should read	... tedious ...
p.65	line 12	... to one-hundred million mathematical operations. Should read ... by many orders of magnitude.		
p.78	line 30	... dynamic ...	should read	... densitometric ...
p. 146	line 29	... minima ...	should read	... minimum ...
p. 146	line 30	... maxima ...	should read	... maximum ...

LIST OF TABLES

CONVENTIONS, NOTATION AND ABBREVIATIONS

In this section the conventions of notation, abbreviations and symbols are described.

Footnotes and endnotes

The dagger symbol (†) indicates an accompanying footnote which appears at the bottom of the same page. Endnotes are only used for the details of manufacturers and suppliers of materials and equipment used in the various investigations. The existence of endnotes is indicated in the text by a superscript numeral in parentheses (for example; Kodak X-Omat(2) film). Endnotes are placed in Appendix C.

Mathematical and statistical conventions

The notation used for descriptive statistics which are also parameters of images or two-dimensional sequences conforms to current convention in image processing (Jain, 1989). Thus,

mean value	is denoted by	μ	rather than by	$\bar{x}$
standard deviation	is denoted by	σ	rather than by	s
correlation coefficient	is denoted by	ρ	rather than by	r

However, where it is necessary to indicate estimates derived from sample averages this is signified by a circumflex overmark, ^. For example, $\hat{\mu}$ indicates an estimated mean.

Continuous images are represented as functions of two independent variables, thus: $f(x,y)$, $g(x,y)$, ... Sampled images are represented as two-dimensional sequences, thus: $x(m,n)$, $x(m,n)$, ... Corresponding functions in the spatial and Fourier domains are denoted by the corresponding lower-case and capital letters, thus: $F(\xi x,\xi y)$ is the Fourier representation of the continuous image $f(x,y)$; and $X(\omega_1,\omega_2)$ is the Fourier representation of the sampled image $x(m,n)$.

SYMBOLS

The following conventional symbols are used:

Symbol	**Meaning**
C(p,q)	cross-correlation function
e	transcendental number (2.71828...)
exp	exponential [$=e^x$]
$\mathscr{F}$	Fourier transform
g(x,y)	two-dimensional x-ray attenuation pattern
H	frequency response
H_0:	null hypothesis
H_m:	*m*th hypothesis
kV	Kilo-volts potential
ln	natural logarithm ($\log_e$)
m,n	horizontal and vertical coordinates of a digital image
μ	mean
μm	micrometres
mA	milliamperes
mm	millimetres
n	number in a sample
ns	not statistically significant
n(x,y)	two-dimensional noise
$\mid N_t \mid^2$	noise (power) spectral density
ξ	continuous spatial frequency variable
ρ	correlation coefficient
$\rho_{1\text{-step}}$	one-step correlation coefficient
ρ_{xy}	correlation coefficient between x and y
p	probability
r	received signal
r_s	Spearman's rank correlation coefficient
R_{sr}	correlation between s and r
R_{ss}	autocorrelation of s

$R(m,n)$	two-dimensional autocorrelation function
σ	standard deviation
σ^2	variance
σ_{xy}	covariance
φ_m	*a priori* probability of correct match for *mth* subimage
Σ	correlation matrix
s	transmitted signal
s_m	*mth* subimage
S_x and S_y	standard deviation for differences of double determination
S^*	complex conjugate of signal, s
U	area under the ROC curve
W	Shapiro-Wilk's statistic
Φ	image matrix
ω	discrete spatial frequency variable

ABBREVIATIONS

The following abbreviations are used:

Abbreviation	**Meaning**
ACF	Autocorrelation Function
ANOVA	Analysis of Variance
APRI	Active Penetrating Radiation Imaging
CCD	Charge-Coupled Device
CI	Confidence Interval
ERF	(Gaussian) Error Function
FFT	Fast Fourier Transform
FWHM	Full-Width at half-Maximum (of the ACF)
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MLP	Maximum Likelihood Processor
MR	Magnetic Resonance

MRF	Markov Random Field
MRF-1	First-order Markov Random Field
MRI	Magnetic Resonance Imaging
PDF	Probability Density Function
PPRI	Passive Penetrating Radiation Imaging
PFA	Probability of False Acquisition
PSR	Peak-to-sidelobe Ratio
ROC	Receiver Operating Characteristic
SD	(sample) standard deviation
SNR	Signal-to-Noise Ratio
$\mathrm{SNR}_{\mathrm{in}}$	Signal-to-Noise Ratio at the input (to the cross-correlator)
$\mathrm{SNR}_{\mathrm{out}}$	Signal-to-Noise Ratio at the output (of the cross-correlator)
SOP	System Operating Point
SSD	Sum of Squared Difference correlation function

CHAPTER
1

INTRODUCTION

1.1 BACKGROUND TO THE STUDY

An understanding of the growth patterns of the human face is important in the optimal timing of orthodontic treatment and for planning complex procedures needed to correct severe congenital defects in the young patient. In addition, knowledge of the variations in the timing, direction and intensity of facial bone growth is important in assessing the effects of different treatments designed to influence or harness growth.

Unfortunately, our understanding of these variations is quite limited principally because of the restrictions imposed on the study of living human subjects.

Methods of examination are limited to those which are non-injurious and this has meant reliance upon non-invasive imaging techniques. The presently available clinical imaging methods produce images which provide mainly static morphological information. To obtain information about dynamic processes, such as growth, it is necessary to record multiple images separated in time to produce a "temporal image sequence".

In order to extract the dynamic information contained in the image sequence and provide a meaningful analysis, the images must first be orientated and aligned relative to each other, in much the same way as the individual frames of a cine film are aligned during projection. This process of orientation and alignment is known as "spatial registration" and is fundamental to the analysis of any image sequence.

Spatial registration is achieved by locating and matching the unchanging regions in each image of the sequence. This presupposes the existence of stable structural details within the tissues being imaged and requires these details to be detectable throughout the sequence. These requirements pose few problems for the analysis of rapid sequences where the images are recorded at sub-second intervals for the study of rapid phenomena (as in *scanning laser ophthalmoscopy* or *dynamic cardiac imaging).* However, many changes of biomedical interest occur at much slower rates taking many months or years to express themselves. For example, benign tumour growth, atherosclerosis, bone remodelling or tooth migration. The registration and subsequent analysis of image sequences depicting slow changes pose problems which are not apparent in the analysis of rapid sequences. This is because of (i) distortions and differences between the images which arise from alterations in the alignment of the subject relative to the imaging system and (ii) variations in the settings or sensitivity of the imaging system over prolonged periods of time. In addition, while the long-term persistence of some structural details occurs in most hard tissues, changes resulting from growth and/or treatment can often obscure and mask them preventing their detection. As a result, the accurate measurement of long-term changes has generally proved impossible without the use of artificially implanted markers to define regions of stability (Björk and Skieller, 1976; Rune, 1980).

Unfortunately, it is impractical to use implanted markers routinely and for many regions of clinical interest the use of such markers is not possible because of the difficulties of surgical access. It would therefore be highly desirable to find a method for accurately registering long-term image sequences, depicting hard tissue changes, without the need for implanted markers.

The purpose of this study was to develop a method to locate naturally occurring structures for use as reference markers for the accurate registration of serial radiographs depicting long-term changes in growing bone. The many physiological and pathological changes which occur within and between bones of the face and cranium and within growing bones generally, are of such wide interest that the development of an accurate implant free registration system would represent a useful advance in biomedical imaging science.

1.2 OUTLINE OF THE THESIS

The research which led to the writing of this thesis began because the radiographic measurement of tooth eruption and macroscopic changes of bone contours, could not be determined without the use of tantalum markers. To gain insight into these problems it was necessary to develop a method for detecting and locating natural reference sites in the radiographically visible structure of a bone. The description of the development, implementation, testing and application of the detection method forms the substance of this thesis.

The thesis is laid out in a series of 12 chapters each of which describes a separate aspect of the study.

Chapter 2 provides a review of the literature relevant to the subject of the thesis. It is divided into four sections. The first section provides an overview of the currently available methods for the clinical imaging of bone together with their advantages and deficiencies. The second section gives a more detailed review of the physical basis of serial radiographic imaging. In the third section, the concepts underlying image registration are discussed; and the final section presents a review of reference markers used for the registration of long-term serial radiographs of bone.

Chapter 3 describes the aims of the study and examines the nature of the problem investigated in this thesis together with possible approaches to its solution.

Chapter 4 surveys the procedures necessary for the manipulation of digital images together with the basic theory behind the processing of two-dimensional digital signals. This chapter provides the theoretical and mathematical framework for understanding much of the material presented later in the thesis.

In chapter 5 the general method of approach to the problem is outlined, and a general theoretical model of the problem is developed. Analysis of the model, using known solutions from signal detection theory, is described and the theoretical structure of the detection method is defined.

Chapter 6 provides a theoretical model of the features which hinder detection - the temporal noise. The chapter is divided into two main sections: in the first section a simple model of the temporal noise is developed from a set of plausible physical properties of the changes which constitute the temporal noise. In the second section, the relevant features of the model are identified and formalised into a mathematically tractable yet simple statistical model.

Chapter 7 describes the materials and methods used in the experimental development, testing and application of the detection system. The calibration and inherent errors of the experimental apparatus are described together with the methods of handling and analysing data in the form of digital images.

In chapter 8 the validity of the theoretical models, developed in chapters 5 and 6, are examined in a series of experiments designed to test the agreement between the models and actual data. The results are discussed and statistically analysed. As a consequence of the findings, modifications to the detection system are proposed.

In Chapter 9 the theoretical form of the detection system, derived in chapter 5, is developed into a practical detection system. The implementation of the detection system as a computer-based image processor is described.

Chapter 10 looks at the performance of the detection system and examines its inherent errors. In the first section a detailed theoretical description of the possible sources of error is provided. In the second section, the actual errors are examined in a series of experiments using serial radiographs of children with implanted markers.

In chapter 11 the computer-based system is applied to the measurement of human incisor eruption from late intra-alveolar to pubertal eruption stages. The purpose of this investigation was to provide an example of how the detection system can be applied to a problem which has so far eluded detailed analysis. The difficulties and errors involved in the application of the processor are described and discussed. New insights into the pattern of incisor eruption and migration are described.

Chapter 12 concludes the body of the thesis with some general considerations on the consequences of this study and provides a summary of the important findings.

CHAPTER
2

REVIEW OF THE LITERATURE

2.1 CLINICAL IMAGING OF GROWING BONE

2.1.1 Introduction

The bones of laboratory animals and specimens of bone removed from human subjects may be imaged in a variety of ways, at all spatial levels from simple low-power microscopy of gross specimens to computer-enhanced back scattered scanning electron microscopy of individual collagen fibres (Revell, 1992). However, in the living human subject imaging is limited to those methods which are non-injurious and sufficiently penetrating to provide information from deep within the body.

Until very recent times the only method for non-invasively imaging living human bone was by the transmission and attenuation of an x-ray beam - radiography and x-ray tomography. However, during the past two decades, advances in digital computer processing and aperture imaging have fostered the development of new techniques based on the transmission, emission and reflection of other forms of radiant energy, together with new approaches to the use of x-ray imaging, such as dynamic spatial reconstruction techniques (Jaffe, 1982) and computed x-ray tomography.

2.1.2 Imaging by penetrating radiation

Two fundamental forms of penetrating radiation imaging can be recognised

(Hunt, 1976): passive penetrating radiation imaging (PPRI); and active penetrating radiation imaging (APRI). In PPRI the object is illuminated by a source of radiation (neutrons, x-rays, gamma-rays) which penetrates the object. Variations in the relative density of the object lead to an uneven attenuation of the beam and the production of a shadow image referred to generically as a radiograph[†]. In APRI the object itself is made the source of the radiation usually by inclusion of a radioactive source within the object. The image is produced by variations in the concentration and spatial position of the radiation source.

These two different kinds of penetrating radiation systems provide different kinds of information. In practice, clinical PPRI systems primarily record morphological information about the size, shape and spatial positions of internal structures, while APRI systems demonstrate physiological function by careful selection of appropriate radiopharmaceutical or carrier chemicals to which the radioisotopes are attached.

The distinction between PPRI and APRI systems is not usually made in this way for medical imaging systems because many of the forms of penetrating radiation are immediately hazardous to living tissue. The range of imaging possibilities is therefore limited. Four different types of penetrating radiation systems are generally recognised in medical imaging: radiography (including conventional tomography), x-ray computed tomography, nuclear medicine images, magnetic resonance imaging (MRI) and ultrasonography.

2.1.3 **Radiography**

Historically, radiography was the first and is still the most widely used form of transmission imaging used to record the morphological structure of bones.

Conventional radiography, tomography and x-ray computed tomography record an image which depends on differences in the attenuation of the x-ray beam produced by different tissues. The attenuation depends on two principal factors: the atomic structure of the tissue; and the length of the path through the tissue. Bone, and other hard tissues which contain hydroxyapatite produce a marked attenuation approximately proportional to the number of calcium atoms in the path of the beam.

[†] The term "radiograph" as employed in the physical sciences, includes all images recorded by x-ray transmission, including classical tomography and x-ray computed tomography.

Having passed through the tissues, the differentially attenuated beam interacts with a detection system which senses and records the x-ray photons to produce a visible image for final viewing.

The apparatus used to produce radiographs consists of two parts: an x-ray source and a detection system. The x-ray source is usually a tungsten target which is bombarded with a high voltage electron beam containing a range of energies. To produce a more nearly monochromatic x-ray beam, requires a source consisting of a radioisotope.

The detection system generally consists of two parts: a detector and recording medium. The most common form of detector is a fluorescent screen which contains a phosphor powder deposited in a binder. Phosphor powders are crystalline materials which contain small amounts of activator atoms in their atomic lattices. The fluorescent screens act as amplifiers for the x-ray signal by producing a shower of photons of visible light from each x-ray photon which reacts with a phosphor crystal. Calcium tungstate was the principal phosphor used in x-ray intensifying screens until the mid-1970's when rare earth oxybromides and oxysulphides were introduced (Buchanan, 1972). The visible photons produced by the phosphor are recorded on a photographic silver halide film held in intimate contact with the florescent screen.

Alternatively the x-rays can be detected without an intervening phosphor by the direct effect on an x-ray sensitive silver halide film. The spatial resolution and general image quality is better with direct exposure but at the expense of much higher radiation exposure. Direct exposure is almost exclusively limited to small intra-oral radiographs. With either method, the silver halide film requires subsequent chemical development to reveal the final image.

An alternative to photographic film recording is provided by xeroradiography where a selenium plate holds a static electric charge proportional to the intensity of the x-ray beam falling on its surface. The image is viewed by dusting the plate with a fine powder which is then transferred to a paper backing sheet (in the same manner as a photocopier). Xeroradiography, has several advantages over conventional photographic recording of the image, most important of which is the edge enhancement effect which provides a subjectively clearer image. Unfortunately, xeroradiography requires approximately twice the radiation exposure as a conventional screen/film system

(Shaw, 1975).

During the past decade increasing use has been made of radiographic systems which employ photoelectronic detection and recording of x-ray images (James et al., 1985). Such techniques are generally referred to as *film-free* radiological systems (Gröndhahl, 1992). Two basic types of sensing device are used: x-ray sensitive video cameras; and x-ray sensitive charge-coupled devices (CCD). In dentistry, the compact nature of CCD's has allowed them to be developed as intra-oral sensors.

In addition to the basic components of the detection system, practical radiographic systems also employ devices to reduce the level of scattered radiation reaching the detector or film. While scattered radiation increases the volume of tissues which is irradiated, it also adversely influences the quality of the final image by reducing the contrast. The amount of scattered radiation is primarily influenced by the x-ray spectrum and the volume of tissue irradiated. The principal method for reducing the scattered radiation reaching the detector is to use a stationary or moveable grid. Such grids consist of fine slats of radiopaque material arranged parallel to the line of the x-ray path between object and film. Only those x-rays passing from the direction of the source are permitted to pass between the slats. The use of a grid provides better image contrast but leads to an increase in x-ray exposure of the subject.

2.1.3.1 *Classical x-ray tomography*

Tomography, sometimes referred to as laminagraphy, is a radiographic technique in which a chosen layer within an object is clearly imaged while structures overlying and beneath this layer are blurred or not imaged at all. Two basic types of tomography may be distinguished: classical tomography and computed tomography.

Classical tomography is achieved by causing the x-ray source and detection system to move in opposite directions during the x-ray exposure. This relative motion maintains the image of a thin layer at a fixed location on the detection system, while the images of all other layers are moving across the detector. Classical tomography excels in the imaging of small high contrast structures such as bone and dental hard tissues (Barrett and Swindell, 1981b).

2.1.3.2 *X-ray computed tomography.*

X-ray computed tomography, like classical tomography, requires an x-ray source and a detection system. However, the detection system consists of multiple photoelectronic detectors which encircle the subject and feed information about the detected x-ray photons to a computer. The computer then calculates an attenuation profile along a series of lines joining the x-ray source to each detector.

Although three-dimensional images can be calculated by the computer, and the object viewed in two- or three-dimensional space, the spatial resolution achievable with computed tomography is limited to about 0.5 to 1 mm (Barrett and Swindell, 1981b; Engelke, et al., 1993). In addition, an effect known as *beam hardening* and image artifacts known as *edge ringing* limit the accuracy of reconstruction of high contrast objects such as bone (Scudder, 1978). Consequently, in the imaging of bone and dental hard tissues, x-ray computed tomography is unable, at present, to match the clarity and distinction achievable by classical tomography.

2.1.4 Nuclear medicine images

These methods are generally not applied to the morphological imaging of bone. They can be used to monitor physiological activity in tissue by selecting an appropriate radioisotope which localises in specific tissues or at sites of specific biochemical activity.

The gamma radiation emitted by the decay of the radioisotope is detected and used to provide two- and three-dimensional reconstructions of the tissue using a computer. These techniques have limited application in the morphological analysis of growing bone because of poor spatial resolution and low geometric fidelity. Nevertheless, recent advances in this form of imaging make this a potentially useful method of detecting metabolic changes in growing bone (Litton et al., 1990; Germano et al., 1990). However, the need to inject or ingest radiopharmaceutical compounds poses an extra hazard which makes these methods inherently unsuitable for use in otherwise healthy subjects.

2.1.5 Magnetic resonance imaging

Magnetic resonance imaging (MRI) is a relatively recent introduction in medical imaging (Mansfield and Maudsley, 1977). Its principal advantage over conventional x-ray imaging is the absence of exposure to ionising radiation. Unlike x-ray imaging techniques MRI does not depend on differences in electron density but depends instead on differences in proton density and tissue magnetic relaxation properties.

A magnetic resonance image is created by placing the subject in a powerful magnetic field which aligns the constituent protons (hydrogen nuclei). A radio-frequency field is then applied which causes the protons to change their orientation. When the radio-frequency field is removed the protons realign in the magnetic field; in doing so they emit the energy absorbed from the radio-frequency field. This emitted energy is detected by an antenna which surrounds the subject. The image is reconstructed using a computer in much the same way as for computed tomography.

MRI is generally considered to be less useful for imaging bone than for imaging soft tissues because cortical bone has few hydrogen nuclei and therefore provides only a small signal. Consequently bone appears black on the MR image. However, increasing use has been made of MRI examination of the temporomandibular joint because MRI is able to reveal bone surface changes and variations in vascularity.

MRI has some advantages over computed tomography in the imaging of bone: fewer artifacts occur from the presence of dense bone and beam hardening is not a problem. Unfortunately, the long scanning time degrades the spatial resolution and the accuracy of the reconstructed image is limited by the capacity of the data handling computer: for example, to double the resolution requires a quadrupling of the input data (Crooks et al., 1984).

2.1.6 **Ultrasound imaging**

This form of imaging depends on the interaction between high frequency sound waves and tissue. This interaction is dominated by reflection and scattering which returns some of the acoustic energy back towards the source where is can be detected by a sensitive microphone. The scattering and reflection depend on a property referred to as *acoustic impedance* which is determined by the mechanical properties of the tissues. Although ultrasound examination has proved useful in bone mineral determination (Felsenberg, et al., 1993), ultrasound provides only poor images of bone because of the distortion and reflection produced at the interface between bone and soft tissue (Barrett and Swindell, 1981b).

2.1.7 Other imaging methods

There are several medical imaging methods which provide only poor spatial resolution and are not generally used to image bone for example, *thermal imaging* and *electrical impedance imaging* (Newell et al., 1988). Although these methods seem entirely unsuited to morphological imaging of hard tissues, recent advances in electrical impedance imaging may make this a clinically useful method of imaging bone (Chen and Paoloni, 1989)

2.2 METHODS FOR SERIAL IMAGING OF GROWING BONE

Many of the methods described in the previous section can be used to provide serial images of growing bone. However, images recorded for quantitative studies must possess good accuracy and spatial resolution. These restrictions limit the useful methods to those of radiography and conventional tomography. Recent improvements in the accuracy and spatial resolution of x-ray computed tomography and MRI make it possible for these methods to be applied to serial imaging of growing bone. However, the cost of these methods and the radiation exposure involved in computed tomography limits their use.

2.2.1 The limitations of conventional radiography.

Despite the use of radiography and tomography for serial studies of bone growth they have several limitations. Foremost is the restricted spatial resolution, especially for low contrast details. Conventional screen/film systems provide a maximum spatial resolution of about 6 line pairs per millimetre. Higher resolution can be achieved at the expense of greater radiation exposure using finer crystals in the intensifying screens or by using plain x-ray film (for example, dental intra-oral radiography) again at the expense of increased radiation exposure.

Secondly, conventional radiography produces a two-dimensional shadow image where depth information is lost. Unfortunately without additional independent evidence it is not possible to recover information about the depth of any given structure.

However, depth information is recoverable to a limited extent if several films are exposed from different perspectives allowing tomomsynthetic reconstruction of the object image (Webber, et al., 1985) or by stereoscopic viewing of two simultaneously exposed films (Rune, 1980).

Where multiple images are not available the lack of information on depth in the object, true three-dimensional quantification and understanding of changes is not possible.

2.2.2 Standardization of image geometry

There are two basic methods for establishing a constant geometric relationship between the images to be registered. The first method, is to standardize the image projection geometry thereby reproducing the same relationship between x-ray source, object and film at each exposure. The second method, is to correct any geometric distortion between the images by performing a non-rigid registration, in which the geometry of one of the images is elastically deformed to force the image coordinates into the same relationship.

2.2.2.1 *Standardization of the image projection geometry.*

There are two basic types of device which can be used to standardize the projection geometry: coupled devices and uncoupled devices. Coupled devices are aiming mechanisms which rigidly couple together the film, subject, and x-ray source while uncoupled devices do not rely on physical contact between the film subject and x-ray source.

Coupled devices.- There are two types of coupled aiming device in common use: occlusal stents; and cephalostats. Occlusal stents for use with intra-oral radiography are aiming devices which attach the teeth to the x-ray cone using a customised bite block which fits the occlusal surfaces of the teeth and holds the film. Cephalostats use the same principal of rigid coupling between the components of the x-ray system but unlike stents which use the teeth to secure the device, a cephalostat uses a framework which hold the subject using ear posts and an infra-orbital or nasal bridge pointer (Broadbent, 1931; Solow and Kreiborg, 1988).

The use of ancillary mechanisms to improve the reproduction of imaging geometry have been proposed both for stent based and cephalostat based systems. These ancillary devices generally consist of a light beam attached to the x-ray source together with a mirror or reflective disc attached to the subject. Although, most of these systems have been designed for research rather than clinical use (for example, Price, 1973; Tan and Van Aken, 1977), clinical systems have been constructed which incorporate these features. For example, Solow and Kreiborg (1988) described a cephalostat unit which incorporated six laser beams projected onto the face and head to aid in the reproducible positioning of subjects with gross asymmetry of the skull.

Uncoupled devices.- Coupled devices suffer from the disadvantage that they must rely on external features to orientate the x-ray source and film. If the external anatomical features change position relative to the underlying bone the image geometry will alter (Tan and Van Aken, 1977). In an attempt to overcome this problem Björk (1968) developed a cephalostat which permitted video feedback prior to exposure of the film. This technique used a cephalostat to provide coarse positioning of the head followed by fine positioning using a low continuous x-ray exposure and an image intensifier to compare the bone image with that from a previous film. Once the position of the bone appeared the same in the two images the film cassette and grids were then rapidly brought into place with a motor-driven mechanism, and the film exposed. This type of system clearly increased the x-ray exposure to the subject and is now only used for detailed studies in patients with major cranio-facial deformities (Solow and Kreiborg, 1988).

2.2.2.2 *Correction of geometric distortion.*

Meaningful correction of geometric distortion depends on the ability to globally warp or transform the coordinates of one or both images. Global warping can be viewed most easily by thinking of one of the images printed on a sheet of rubber. Registration is achieved by deforming this rubber sheet until all points on the sheet correspond exactly with those on the other, undistorted image. However, for such elastic deformation to be effective the structures in the image must either occupy a very limited range of depths or the images must be recorded from exactly the same perspective. This is because variations in projection between exposure of the radiographs leads to alterations in the superimposition of structures at different depths in the bone (Figure 1.1).

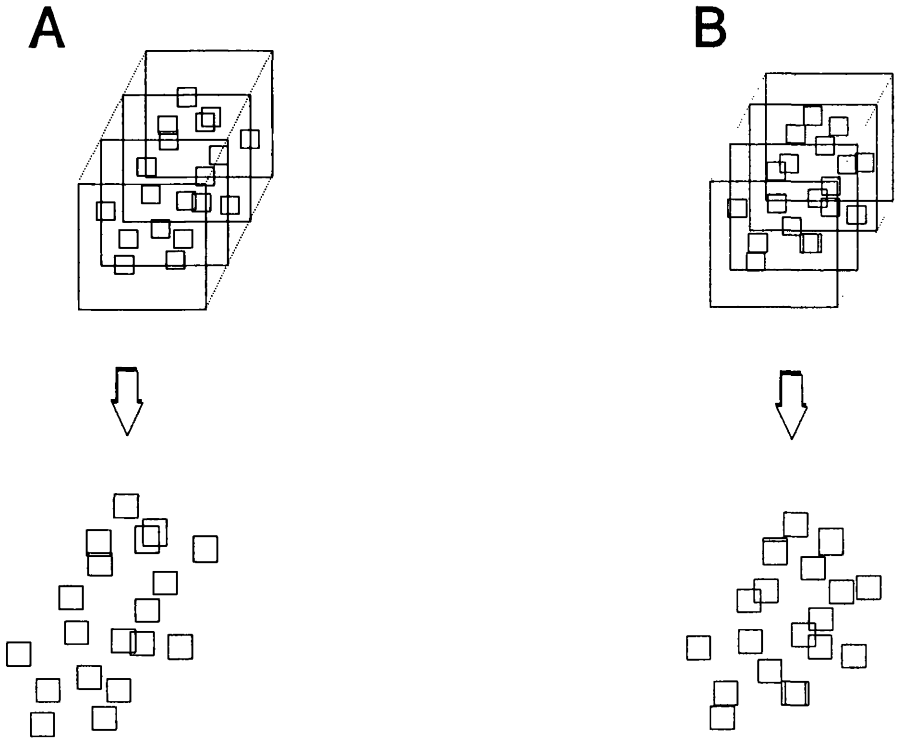


Figure 1.1 Diagrammatic representation of the changes induced in a 2-D image by small differences in imaging geometry. A and B represent the same object with four planes of structure.

Correction of this type of geometric distortion requires a different deformation for each separate layer in the object. This is not usually possible without additional information about the depths of structures within the bone. However, Dunn and van der Stelt (1992), in an initial investigation of a new model of imaging geometry, have shown that correspondence matching, employing projective invariants, may overcome the problems of registering unstandardised serial intra-oral radiographs.

However, where the image geometry is well controlled usually below 3 degrees the geometric distortion can be approximated by a simple affine distortion model (that is, a simple shift or rotation of the image coordinates to simulate a rotation out of the film plane). In this case the planar component of the geometric distortion can be corrected using an affine matrix transform (Jeffcoat et al., 1984).

2.3 SPATIAL REGISTRATION IN THE ANALYSIS OF IMAGES

2.3.1 Introduction

Whenever it is required to detect changes or to integrate data from two or more images, for the results to be meaningful the images must be spatially registered (Barnea and Silverman, 1972). Spatial image registration is defined as the process of geometrically aligning or overlaying two or more images of the same "scene". However, the process of registration does not necessarily require the images to be physically aligned or overlaid; it is often only necessary that the coordinate data defining the scene are accurately known for each image, thereby allowing a "mathematical registration" of the data.

Spatial image registration is an essential stage in the analysis of all forms of "multiple imagery" (Green et al., 1980). Two broad classes of multiple imagery can be recognised in biomedical use: multimodality imagery and image sequences.

2.3.1.1 *Multimodality imagery*

The term "multimodality" refers to a combination of different types of images, usually formed by different types of radiant energy interacting with the tissues under investigation. This form of analysis is becoming an increasingly important part of medical imaging and is used, for example, in planning radiotherapy employing a combination of magnetic resonance and gamma-ray images. Image registration in the analysis of multiple imagery essentially allows for data integration. Registration for the integration of different forms of image requires complex computer based manipulation to apply a "geometric warp" or "elastic deformation" to one of the images before merging the data from both images.

2.3.1.2 *Image sequences*

An image sequence consists of an ordered collection of pictures of a single scene. Image sequences are generally recorded by a single method and are used to gather data which it would be difficult or impossible to obtain from a single image. Two fundamental types of image sequence are recognised: temporal; and spatial, depending upon whether the images forming the sequence are primarily separated in time or in space.

The analysis of temporal image sequences is particularly important in bio-medical investigations where it is frequently necessary to measure images taken at different times to gather information on changes caused by physiological or pathological processes.

2.3.2 **The basis of spatial registration**

Much of the theoretical background to spatial image registration has been derived during the past two decades as a result of the development of automatic image processors for the analysis of the vast quantity of imagery returned to earth from satellites and spacecraft (Green et al., 1980).

While human registration of images is essentially limited to manual registration of rigid two-dimensional images (for example, conventional radiographs) modern computer-based systems can register geometrically distorted images and large data sets with a level of precision unobtainable by a human operator (Ton and Jain, 1989). However, the basic stages which constitute registration are largely the same whether performed by human or machine.

The process of spatial registration is based on the identification of reference points which precisely locate corresponding pairs of image coordinates (Barnea and Silverman, 1972). The registration procedure then involves modifying the geometry of one image so that the network of reference points is made to have the same spatial relationship in both images.

2.4 REFERENCE MARKERS FOR THE REGISTRATION OF RADIOGRAPHS OF GROWING BONE.

Reference points are derived from objects or structures which are common to the images to be registered. In the case of living tissue the structures which provide reference points are generally referred to as *reference markers*. There are two forms of reference markers which can be used for the registration of long term radiographic sequences of bone: implant markers; and natural markers.

2.4.1 Implant markers

Implant markers are small inert objects, usually fabricated as pins or spheres, inserted into the tissue which act as easily identifiable reference points. Where radiographic imaging is to be used the implants must possess a radiopacity greater than the surrounding tissue otherwise the implant will need to be excessively large in order to be detectable. For this reason, implant markers for bone are usually metal with a high atomic number. The inertness of the implant is important for two reasons: first, any reaction whether physical, chemical or biological with the tissues into which they are inserted will lead to an unnatural alteration of the tissue and possible harm to the subject of the investigation; and secondly, reaction with the tissues may lead to dissolution, loss or movement of the implant.

To provide the required combination of radiopacity and inertness, gold, silver, chrome-cobalt and chemically hardened tantalum have been used to fabricate implants.

2.4.2 Historical review of implant markers in radiographic studies of bone

Implant markers, of various types, have been used in studies of bone growth for over two hundred years (Duhamel, 1742; Hunter, 1770). They were first employed as reference markers for the radiographic measurement of bone growth by Dubreuil (1913) who implanted metallic pins in the leg bones of rabbits. However, it was only following the publication by Bisgard and Bisgard (1935) of their classical study demonstrating the relative positional stability of paired metallic markers in the bones of experimental animals that implant markers began to be widely employed for the accurate radiographic measurement of bone growth.

Subsequently, the deliberate pairing of multiple metallic implants was adopted in other radiographic investigations of bone growth in animals. For example, Haas (1938), used multiple silver implants together with plain frontal and lateral abdominal radiographs to demonstrate that elongation of canine lumbar vertebrae occurred at the epiphyseal plate and not by interstitial growth of the vertebral body.

The radiographic-implant technique was adapted for use in humans by Gill and Abbott (1942) who developed a closed surgical technique for implanting single 0.8 mm hollow "vitallium" needles into the femur. Although, they discussed at length the work of Bisgard and Bisgard (1935), the authors chose to use single implants in each bone to make accurate measurements of the rate of growth of the long bones in children with asymmetrical increase in leg length.

By the late 1940's evidence from studies of growth in primates using vital staining techniques began to cast doubt on the validity of natural reference points in the cranial and facial bones (Brodie, 1949). This prompted attempts to combine implant markers with cephalometric radiography for the study of facial growth. Both Hulen (1948) and Black (1948) used gold foil implants to investigate the growth of the mandible and palate in the dog. The technique was developed and refined in animal studies during the early 1950's by Sarnat and co-workers (Gans and Sarnat 1951; Sarnat and Gans, 1952).

Later, the combined cephalometric-implant technique was adapted for use in humans by Björk (1955) who developed a closed surgical technique to place metallic implants into the cortices of the mandible and maxillae. This technique differed from that used by Gill and Abbott (1942) in that implants were solid pins driven into place with a small piston and surgical mallet, thus allowing them to be placed completely beneath the periosteum (Björk, 1963). This was an important advance based on the findings of Lacroix (1948) who had used a combined radiographic/implant technique to demonstrate the migration of the periosteum over the shaft of the metacarpal of a calf. Implants which protruded through the periosteum were prone to gradual displacement resulting from periosteal migration.

Björk's implants, like those of Gill and Abbott (1942), were initially made from vitallium - a chrome-cobalt alloy. However, the poor radiopacity of vitallium and the possibility that chemical reaction within the bone caused "electrolytic drift" of the pins, prompted a change to chemically hardened surgical tantalum (Björk, 1963). These pins were less affected by reaction with tissue fluids and had greater radiopacity and, consequently, could be made smaller while still allowing them to be detected in the radiographic image. Although tantalum markers have become the standard implants for use in radiographic studies, silver (Isaacson and Murphy, 1964) and gold (Mitchell et al., 1975) have also been employed in human studies. Following Björk's pioneering work, the technique of combined cephalometric radiography and metal implants was adopted by several investigators who wished to analyse facial changes in humans and in non-human primates (for example: Reidel, 1972).

During the 1960's and 1970's the greatest impact of the combined radiographic-implant technique was in the study of human facial growth and orthodontic treatment of minor deformities (Björk, 1963, 1968; Björk and Skieller, 1976; Mathews and Ware, 1978). Attempts to apply the technique to the study of growth and treatment effects in subjects with marked facial deformities proved less satisfactory because a high percentage of implants were dislocated or lost (Robertson and Hilton, 1971; Friede et al., 1977; Shaw, 1977).

The use of implants to provide reference points at bony sites outside the face was uncommon until the development of the *roentgen-stereophotogrammetric technique* by Selvik (1974) which prompted several kinematic studies of human joints (Egund et al., 1978; Van Dijk et al., 1979) and studies of tumour growth in extra-skeletal sites (Claesson et al., 1978).

Since the late 1970's ethical considerations have generally prevented the use of implants in human growth studies (Mathews and Ware, 1978; Poswillo, 1984) and concern over x-ray exposure has limited the radiographic-implant technique in humans to diagnostic studies of gross facial deformity (Rune, 1980). However, the ethical limitations imposed on human studies have not applied to animal investigations where the implant technique has been refined to provide stable radiographic markers for dynamic studies of cartilaginous and fibrous structures as well as bone (Laitman and Crelin, 1980; Koskinen-Moffett et al., 1981).

2.4.3 Validity of implanted markers as reference points for spatial registration

The usefulness of implanted markers as reference points for spatial registration depends on two basic qualities. First, they must be a valid representation of the tissue in which they are implanted; and secondly, they must be uniquely identifiable in the images of the sequence.

Validity requires that the implants remain immobile in relation to the bone during the period over which the radiographic sequence is recorded. That implants are stable relative to the bone in which they are placed has generally been accepted for over two centuries and was founded upon the perceived maintenance of the relative positions of pairs of markers in the same growing bone (Duhamel, 1742; Humphry, 1864). Although seldom stated explicitly, the stability of the positional interrelationships of pairs of implants was used to infer the stability of the implants to the bone by the following argument: either both implants had moved, relative to the bone, by the same amount, in the same direction and over the same period of time; or both implants were immobile in relation to the bone in which they were implanted. The former explanation is seen as highly improbable and therefore the later explanation is accepted.

The evidence from early gross anatomical studies was supported by later radiographic studies (for example, Bisgard and Bisgard, 1935). However, it is important to note that, in these later investigations, paired implants were employed not as a test of implant stability but to refute the notion of interstitial bone growth.

While metallic implants remain grossly stable in the dense cortices of long bones over short periods of time, it is now clear that in the finer cortices of the facial bones implants frequently drift and are occasionally lost from the bone (Björk, 1963, 1968; Reidel, 1972). This hazard appears to be particularly common where surgical access is limited for the placement of implants and also where bone growth is followed over many years (Reidel, 1972; Mathews and Ware, 1978). It is necessary, therefore, to *test* the stability of implants as each successive image is recorded. This is done by examining the relative positions of several implants in subsequent images and comparing them to the first image.

However, Rune (1980), using three-dimensional radiographic stereophotogrammetry has shown that there is a "settling-in" period, often lasting many months, during which the implants show a mean variation of approximately 0.25 mm (standard error $\approx$ 0.1 mm). Following this initial period, Rune (1980) reported that implants were stable within the technical error of the recording method ($\pm$ 50μm.) so long as remodelling did not uncover the implant. Consequently, comparison of the positions in an initial image recorded at, or soon after, the time of implant placement may not provide the most appropriate evidence on which to establish later stability.

The second attribute of an implant, being uniquely identifiable, is provided by their high radiopacity relative to bone. However, this is not always sufficient where several implants are used in the same site or in overlying anatomical regions (Björk and Skieller, 1976; Rune, 1980). Under these circumstances it is often difficult to determine corresponding implants in serial images. This problem has been partially solved by placing implants of different sizes at different anatomical locations (Björk and Skieller, 1976) or, in the case of stereo-radiographs, by calculating a disparity matrix for all possible matching pairs of implants and selecting the pair with consistent minimum disparity (Rune, 1980). Nonetheless, uniquely identifying matching implant markers can pose problems, for example in the presence of metallic dental restorations or where implants are placed at corresponding sites in the two maxillae (Buck and Hodge, 1975; Björk and Skieller, 1976).

2.4.4 **Natural markers.**

Natural markers are reference points supplied by structural elements of the tissue being imaged. Unlike implanted markers, the anatomical sites at which the natural markers are found cannot be determined beforehand; and it is not usually possible to verify the stability of natural markers on an individual basis. However, unlike implanted markers, they do not require surgical intervention for their placement nor do they require the introduction of foreign material into the tissues.

Since the mid-1920's, supposedly stable, natural markers of one form or another have been used to provide quantitative estimates of bone growth. Three principal types of natural markers may be distinguished:

(i) geometrically defined lines and planes;
(ii) anatomically defined structures and natural boundaries;
(iii) bone microstructure (persistent trabecular patterning).

2.4.4.1 *Geometrically defined lines and planes.*

This class of reference structures is almost exclusively limited to the face and skull. They arose, during the 1920's and early 1930's, from attempts to extend classical anthropometric and craniometric studies to the analysis of radiographs of the skull (Broadbent, 1931). Their use is limited to registration of serial lateral cephalometric radiographs, although attempts have been made to use them for registration of rotational tomograms and oblique lateral jaw radiographs (Feaseby, 1981). However, the strict definition of many of the points used to define these planes and lines requires a standard lateral projection of the skull.

These reference markers have seldom been subject to any objective check on their stability, nonetheless, a voluminous literature has developed concerning cephalometric reference sites (Krogman and Sassuoni, 1957). By the late 1940's evidence from studies of growth in primates using vital staining techniques began to cast doubt on the validity of measuring sutural growth in the face by radiographic registration using cephalometric planes (Brodie, 1949). Further evidence that cephalometric planes alter during growth was provided by the early implant studies of facial growth (Björk, 1955; 1963).

The points which define the ends of the most commonly used cephalometric planes, have been shown to lie on anatomical contours which are subject to continuous remodelling during growth (Björk, 1969; Björk and Skieller, 1977). It is now considered inappropriate to use such geometrically defined lines and planes for quantitative analysis during periods of active growth (Björk, 1969; Feaseby, 1981; Björk and Skieller, 1983).

Although the use of cephalometric planes as registration sites for quantitative studies is largely historical, they continue in widespread clinical use because of their ease of application and high reproducibility compared to other, potentially more valid, methods (Cook and Gravely, 1988). However, this method of registration is generally considered to be valid for studies of tooth movement in adults, where the magnitude of the tooth movement is much greater than the bony remodelling at the cephalometric points (Cook and Southall, 1989).

2.4.4.2 *Anatomically defined structures and natural boundaries*

This class of reference marker is historically the oldest and most widely used. The anatomical structures used for registration vary depending on the bone being radiographed.

The Long bones.- The principal anatomical markers in the long bones are:

i) nutrient canals as they pass through the cortex;

and ii) transverse lines ("Harris lines")

The use of nutrient canals as reference markers was first reported by Digby (1915) who used them to estimate diaphyseal growth rates of the arm and leg bones of adult humans. The validity of the nutrient canals as stable reference markers was first questioned by Payton (1932) who observed remodelling at the periosteal and endosteal margins of the canals using chemical bone markers in young pigs and by Bisgard and Bisgard (1935) using metallic implants in the long bones of goats. The instability of nutrient canals was later confirmed in human subjects by Gill and Abbott (1942) using chrome-cobalt implant markers inserted into the humerus and femur. However, Gill and Abbott (1942) and Lee (1968) claimed that nutrient canals demonstrated reasonable stability for clinical measurements.

Although Harris (1926) was the first to report the use of, what latter became known as, "Harris lines" or "Harris lines of arrested growth", persistent transverse lines seen in radiographs of the long bones had been reported earlier by Phemister (1918) and by Stettner (1921). Park et al., (1931) demonstrated the formation of these lines following acute lead poisoning in children and postulated their origin as a densely packed lattice of trabeculae. Sontag (1938) demonstrated similar lines or stria in the tarsal bones of neonates attributing them to the change from prenatal to postnatal environment. Garn et al., (1971) demonstrated the widespread nature of these lines and compared them to the neonatal lines observed in histological sections of dentine and enamel.

Despite the apparent widespread disposition of transverse lines in the bones of the hands, feet, arms and legs, the unpredictable timing of their appearance and generally poor radiographic contrast has limited their use as reference markers.

The validity of persistent transverse lines as stable reference markers in humans was originally suggested by Harris (1926) who demonstrated an apparently constant inter-relationship between pairs of lines in the same bone over periods up to 8 years. This was supported by Gill and Abbott (1942) using measurements of the relative positions of the lines and implant markers inserted in the bony cortex. However, these studies examined the lines at a gross scale and without the aid of constant projection geometry. Consequently, these studies provide only limited evidence of the stability of transverse lines despite the widespread acceptance of their validity as reference markers (for example, Roche, 1965).

Attempts to provide a more objective assessment of the stability of transverse lines were first undertaken by Garn and co-workers (Garn et al., 1968; Israel, 1968). Colbert (cited by Garn et al., 1968) compared the structural stability of paired lines seen in serial radiographs over a mean period of 10 years in four individuals using a scanning micro-densitometer. All the lines examined showed a decline in mineral content but the loss of mineral was not always uniform. That is, some lines moved relative to their former position. However, the movement was small (0.2 - 0.5 mm). In the same study Garn et al., (1968) examined the positional interrelationship of pairs of transverse lines in serial radiographs of the tibia of 28 boys and 26 girls over a mean period of 4.1 years (range 1-11 years) and 5.6 years (range 1-17 years) respectively. The distances between the pairs of lines was measured to the nearest millimetre and all values were corrected for magnification. The mean change for the two groups were 0.4 mm (SD 1.5 mm) and 0.7 mm (SD 2.01 mm) respectively for boys and girls. Although the authors concluded that the lines were stable, no method errors were given and consequently, it is not possible to assess whether the individual variability represented random measurement error or true instability.

Jaws and cranium.- Three principal types of anatomical markers have been suggested for registration of jaw radiographs. These are:

i) neurovascular canals or foraminae;
ii) periosteal and endosteal surface contours;
iii) developing and ankylosed teeth;

In the mandible there are eight main anatomical reference markers (Figure 2.2) while in the maxilla there are only three main anatomical reference markers (Figure 2.3).

The validity of anatomical markers in the jaws as stable reference points has been examined by serial comparison with implant markers. However, the majority of studies have not formally examined these relationships, relying instead on general observations.

Such studies include those Björk (1963, 1969), Björk and Skieller (1976, 1977, 1983) and Mathews and Payne (1980). A notable exception was the study by Julius (1972) who examined the relationship between implant markers and anatomical structures in the maxilla and mandible in a group of 21 children over an average period of 3.2 years. He concluded that the mandibular structures described by Björk (1969) were stable with the exception of the third molar germ where teeth had been extracted or where orthodontic space closure had occurred. In the maxilla, the only structure which exhibited a constant relationship to the implant markers was the infra-orbital foramen.

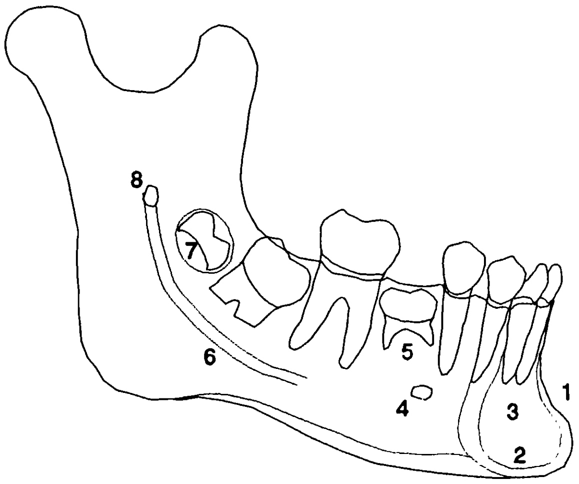


Figure 2.2 Anatomical reference markers in the mandible.

Key: 1 chin contour above pogonion; 2 inferior endosteal contour of the symphysis; 3 prominent trabeculae; 4 mental foramen; 5 ankylosed tooth; 6 mandibular canal; 7 inferior aspect of developing molar; 8 mandibular foramen.

Although, in general, many of these structures may be stable within the limits of the spatial resolution of the radiographs, serious doubt has been cast on the usefulness of many of these anatomical reference sites. For example, it has been reported that orthodontic space closure (Julius, 1972), extra-oral traction (Björk and Skieller, 1976), severe mandibular rotation (Björk, 1963) and fibro-osseous bone disease (Goldberg and Sperling, 1981) may all affect the stability of such reference markers.

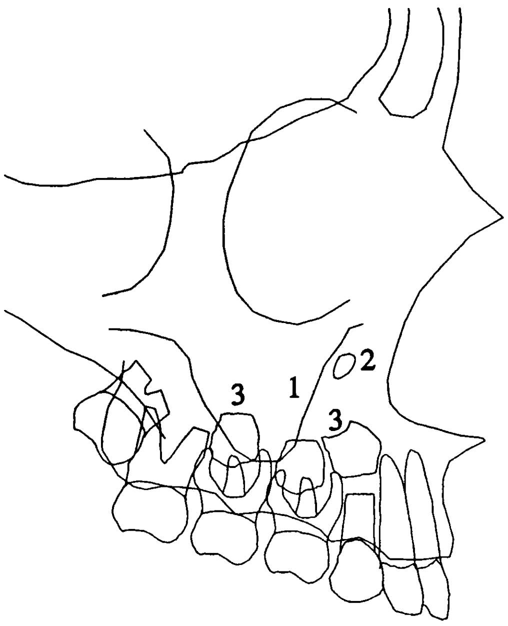


Figure 2.3 Anatomical reference markers in the maxilla: 1 anterior contour of the zygomatic buttress; 2 infra-orbital foramen; 3 apical contour of developing teeth.

Numerous anatomical reference markers have been suggested for the registration of cranial radiographs. However, implant markers have not been employed in the cranial base of man and consequently the validity of these anatomical markers cannot be examined in the same way as for the jaws. Nevertheless, the serial radiographic morphological studies of De Coster (1951), Björk (1963) and Björk and Skieller (1983) and the cross-sectional histological studies of Melsen (1974) indicate that several of the bony contours associated with the anterior cranial base are generally stable after the age of 7 years although slow and unpredictable remodelling of these contours can occur.

2.4.4.5 *Bone microstructure (persistent trabecular patterning)*

Bone microstructure is a term devised by Israel (1968) to describe the fine bony detail seen on radiographs in contradistinction to macrostructure such as Harris lines. Microstructure, as seen on conventional radiographs, represents the individual trabeculae within the medullary cavity which abut against the endosteal margin of the bone (Paul and Trotts, 1964).

Trabecular details have been suggested and employed as *in vivo* bone markers in several detailed radiographic studies (Roche, 1965; Poultón, 1968; Ackerman, 1976; Björk and Skieller, 1983). Although the studies of Poulton (1968), and Björk and Skieller (1983) have shown that the more prominent elements of the microstructure have a constant relationship to implants markers in the same bone it appears that in only one radiographic study have deliberate attempts been made to verify the stability of fine bone microstructure. Israel (1968) investigated the stability of the smallest visible bone trabeculae, in serial periapical radiographs taken over a fourteen year period in three adults. This was done in order to establish their validity as *in vivo* bone markers for the measurement of such changes as tooth movement, tooth wear and alveolar bone loss. Linear sections of corresponding radiographic image detail were recorded in the form of densitometric traces using a Joyce-Loebl scanning microdensitometer. The correspondence between traces and the stability of trabecular details was initially assessed by eye followed by a more objective evaluation using graphical plots of the cross-covariance between traces. From these analyses the long term persistence of many individual trabeculae was demonstrated. However, the three subjects used in this study were all adults where only minimal change might have been expected in the bone structure.

Further supporting evidence for the validity of registration using bone microstructure comes from studies employing photographic subtraction of serial periapical radiographs. Under these circumstances bone microstructure is more commonly referred to as anatomical noise (Gröndahl, 1987). In this method of registration the radiographs are generally aligned by searching for the point at which there is maximal cancellation of stable detail between two serial films where one of the films is tonally reversed. The registration of serial radiographs by subtraction of the bone microstructure has proved highly successful in the evaluation of periodontal bone loss (Gröndahl and Gröndahl, 1983; Hausmann et al., 1985) and in detailed studies of root resorption and bone repair (Price, 1973).

CHAPTER
3

THE AIMS OF THE STUDY: THE PROBLEM AND POSSIBLE APPROACHES TO ITS SOLUTION

3.1 INTRODUCTION

The basis of this investigation was to develop a method which would allow the accurate spatial registration of two or more serial radiographs depicting a growing bone, using only the information available within the radiographs.

3.1.1 The conceptual basis of the problem.

Two or more images are said to be spatially registered if their respective coordinate systems are the same with respect to some fixed origin. In the absence of external information the only way to judge whether two images occupy the same coordinate system is by measuring the similarity of the images. If the images share no common detail it is not possible to register them. However, if they share some common detail and this detail can be unambiguously located then an arbitrary, but fixed origin can be established.

Therefore, *the essential feature of this problem is the ability to detect the presence, and relative location, of common reference points within the images.*

3.1.2 The existence of "stable" structure in living bone

With respect to radiographs of bone, this presupposes the existence of spatially stable structures within the bone which are unlikely to exist. However, provided the level of instability is much smaller than the resolution of the images then such an approach is feasible.

With this concept of "stability" it is clear that the ultimate limiting factor in determining if the images occupy the same coordinate system is the spatial resolution of the images.

In the case of conventional clinical radiographs recorded with a screen/film system, total movement or instability less than about 0.1 mm can be regarded as stability for all practical purposes. While acknowledging that topological similarity does not indicate stability, there is ample evidence from growth studies employing implant markers, that some structures within bone remain stable at this level of accuracy for periods of several years. For example: Poulton (1968) and Björk and Skieller (1983) have demonstrated the long term stability of well defined radiopaque details in medullary bone of the mandible using serial tomography and serial radiography. Furthermore, Rune (1980) observed that following an initial "settling phase", tantalum implant markers remain in stable configurations, within the technical error of the recording method ($\pm 50\mu m$), over periods up to 4.6 years in the maxilla and frontal bone, and up to 3.3 years in the mandible in children, thereby, implying stability of the surrounding cortical bone.

3.2 POSSIBLE APPROACHES TO SOLVING THE PROBLEM

3.2.1 Important facts about the problem

1). The images are clinical radiographs, recorded using a screen/film system giving images with poor radiographic contrast and low spatial resolution but large dynamic range. In addition, the different images are likely to be recorded with different brightness and contrast.

2). Large scale features, which typically have the greatest contrast, are unlikely to be stable despite their similar appearance in successive views. Registration must therefore generally rely on small scale features which form part of a homogeneous texture, often with poor contrast.

3). The radiographs will generally be recorded from similar, although not identical perspectives. However, absolute control of the projection geometry is not possible leading to slight, but unknown, rotational differences in the bone relative to the x-ray source and film. Because bones tend to have large depths relative to their absolute size, even minor rotations lead to variable superimposition of structures located at different depths within the bone.

4). The images are recorded at sufficiently large intervals that much of the internal structure of the bone will have changed as a result of remodelling and restructuring.

5). For radiographs recorded with positioning devices (such as a cephalostat) the approximate rotational orientation of the images will be known to within ± 4° but the relative translational positions of the serial images can often only be estimated to within ± 4 mm (Baumrind et al., 1992).

3.2.2 Possible approaches to the problem

Reference points may be located in images by manual, semi-automatic or automatic methods. Manual approaches to the detection of stable structures, by accurately matching details between radiographs, have clearly failed to provide reference sites with either the level of confidence or the precision provided by implanted markers. While the human visual system can detect likely sites of stable structure it cannot provide the level of objectivity required for this task.

Machine detection and matching can offer objectivity and provides two clear advantages over human detection: precision; and the ability to accurately quantify the match.

The detection and confirmation of a correct match requires the calculation of a measure of similarity between the images and the location of the translational offset coordinates of the point of peak similarity, which is assumed to be the point of correct registration (Barnea and Silverman, 1972; Pratt, 1991). Unfortunately, this approach is inherently sensitive to differences between the two images to be matched, which is a particular difficulty in the present problem. In addition, because major bony contours and large scale features are likely to be subject to continuous remodelling, reliance must be placed on the smaller "texture" feature of the bone with inherently low contrast which generally increases the sensitivity to noise and hinders the measurement of similarity.

This sensitivity to noise and differences between the images can be reduced by the systematic modification of the image data and systematic selection of an optimal similarity function tailored to the particular noise encountered in the images.

The registration methods derived by a systematic approach are generally referred to as *optimal registration processors* or *registration filters* (McGillem and Svedlow, 1977). The concepts underlying the development of optimal registration processors were originally adapted from the theory of military radar and sonar target location, where the displacement between the images to be registered is viewed as analogous to the time delay between transmission and return of the radar or sonar signal (McGillem and Svedlow, 1976, 1977).

There are two obvious advantages to the systematic design of a registration method: first, the method of registration can be tailored specifically to the given set of images and noise sources thereby providing an optimal solution to the specific registration problem; second, if a model-based approach is used to formulate the registration method, theoretical analysis of the model will provide an indication of the limits to the accuracy of the registration method.

Because of these advantages and the need to perform the detection of small, low contrast features in the presence of substantial noise, the systematic design of an automatic registration processor appears, in principle, to offer a solution to the problem and forms the basis of the approach adopted in this study.

CHAPTER
4

SIGNAL AND IMAGE PROCESSING

4.1 INTRODUCTION

The theoretical and experimental investigations which form this study rely on the ability to manipulate or process data, in the form of two-dimensional images, and thereby extract information which would not otherwise be available. The basis of this approach is founded in the branch of science known as signal processing.

A signal can be defined as "a function that conveys information, generally about the state or behaviour of a physical system" (Oppenheim and Shafer, 1989). Signals occur in two basic forms: continuous or discrete. Continuous signals, sometimes known as analogue signals, vary continuously as a function of time or space. An example of a continuous signal is the voltage produced by a microphone in response to sound. On the other hand, a discrete signal is one which takes only a finite number of values or exists only at discrete points in space or time. An example of a discrete signal is a coded message produced by a flashing lamp. The signal can have only one of two values; on or off. The two types of signal are shown diagrammatically in Figure 4.1.

The processing of signals may be performed by several methods but most processing and analysis is accomplished electronically by computer. However, it is usually very difficult to perform precise operations on continuous signals. Consequently, continuous signals are usually converted to discrete signals which are then processed on a digital computer.

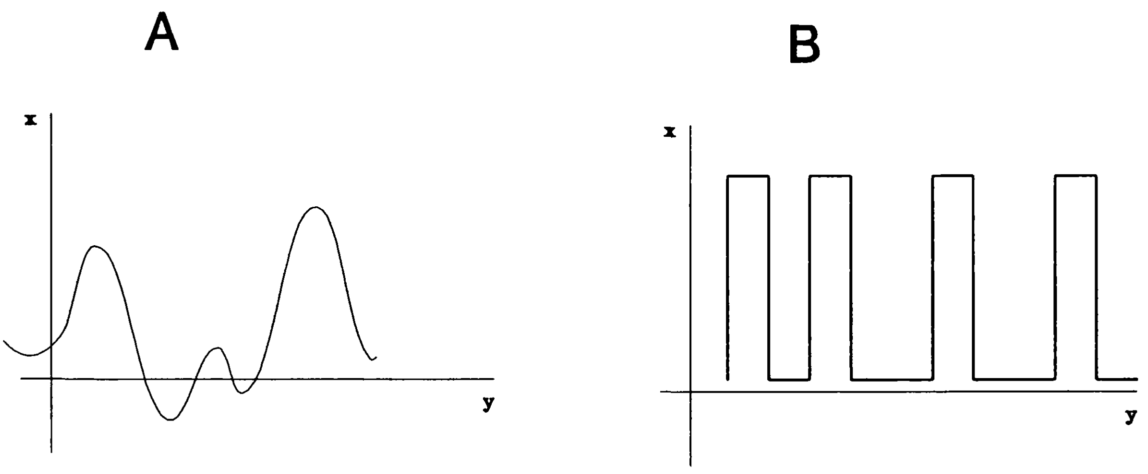


Figure 4.1 Examples of (A) continuous and (B) discrete signals.

Signals occur in a wide variety of forms but a particularly important class of signals are those which occur in a two-dimensional format. In general, any two-dimensional function that bears information can be regarded as an image (Jain, 1989). When considered in this way an image is no longer simply the familiar photographic print or transparency, but is also a two-dimensional array of data. Such an ordered collection of data might represent, for example, the radar cross-section of a military target (radar imaging), the gravitational field over the earth's surface (geophysical imaging) or the x-ray attenuation of a region of living tissue (radiographic imaging). In all cases the image carries information which is contained in some pattern of variations across the array.

4.2 THE DISCRETE TWO-DIMENSIONAL SIGNAL - THE DIGITAL IMAGE.

Digital images consist of discrete picture elements (Figure 4.2), usually abbreviated to pixels or pels, derived by sampling a continuous image. Associated with each pixel is a number which represents the value of the image function at that point. For example, this may be the brightness of a scene, as in a simple photograph, or the x-ray attenuation as in a digital radiograph. The size of the pixels determines the detail

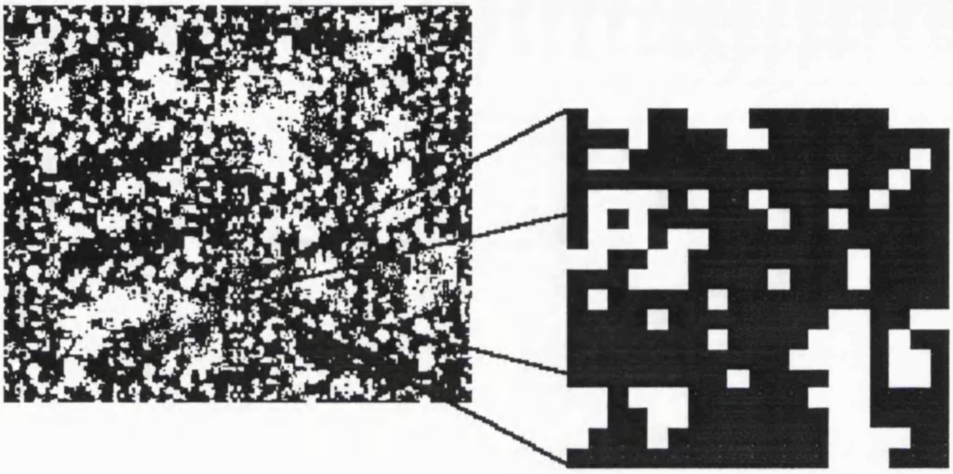


Figure 4.2 A digital image showing the individual pixels in a magnified portion of the image.

visible in the image or more precisely, the spatial resolution. The data, which the pixels represent, may be acquired directly from a digital scanner, as in the case of television pictures, or they may be derived from analogue images, such as radiographs, that have been digitised in the laboratory. The process of acquiring the digital image is often referred to as *image capture*.

4.2.1 Spatial digitization and quantization

Digitization is the name given to the analogue to digital conversion necessary to produce a digital image. It involves two distinct procedures: spatial digitization in which the image is sampled over the region of each pixel; and quantization where the energy levels (or "brightness") of each pixel is given one of a set of discrete levels (commonly 16, 64 or 256 *grey-levels*).

The conversion of a continuous image to a digital image suitable for processing by digital computer is accomplished by sampling the continuous image at several discrete points. The most common method of sampling is to scan the image in horizontal rows and determine the brightness at each sample point. This may be achieved in many different ways dependent upon the speed, cost and accuracy required. For most image processing operations sampling is performed by a television camera with a vidicon tube or image dissector, where the image is converted to an electron image on a photosensitive plate called the target. A fine electron beam scans the target and generates a current proportional to the light falling on the target. This form of sampling is referred to as *scan-in digitizing*. The spatial resolution and quantization accuracy is limited with this method. Higher spatial resolution can be achieved with a scan-out digitizer where the object or continuous image is scanned directly with a fine collimated beam of laser light. The greatest accuracy may be achieved with a flat-bed scanning microdensitometer which employs a high precision optical system coupled to a mechanical scanning arm.

4.2.2 **Two-dimensional sampling theory**

The conversion of a continuous, analogue image to a discrete, digital format appears to imply a loss of information, because the samples are taken only at discrete points. In theory, this is true but only for images with infinite resolution. All real images, however, have finite resolution or more formally, they are *band-limited*. The sampling frequency necessary to recover all the information contained in a continuous band-limited image, is known as the Nyquist rate or Nyquist frequency (Dainty and Shaw, 1974). If the image is sampled below this rate ("under sampled") not only is information irretrievably lost but errors are introduced into the sampled image. These errors are known as *aliasing* (Pratt, 1991) and are not usually detectable except where the original image consists of a periodic pattern. Thus aliasing is occasionally seen as the offensive beating or vibrating pattern on a stripped shirt of a television newscaster.

4.2.3 **Digital image storage**

The binary numbers derived from sampling the original analogue image or scene can be stored in an electronic memory for subsequent analysis and display. The initial electronic memory is referred to as a *frame-buffer* or *framestore*. Because digital images are two-dimensional arrays or matrices of pixels whose quantized levels are powers of 2, it is convenient to arrange the framestore memory into separate planes which correspond to these levels. Each plane in the memory is formed by an array of binary digits (bits) and corresponds to a slice through the image one bit deep. These planes contain only binary information (each pixel can have only two values, 0 or 1) and are

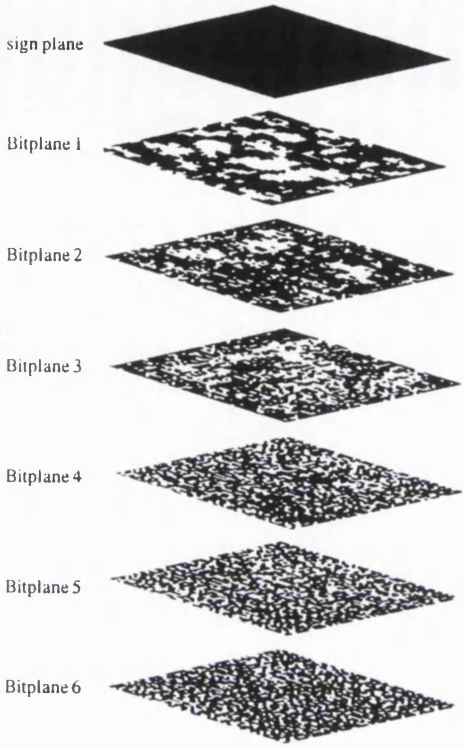


Figure 4.3 3-dimensional diagrammatic view of bitplane image storage showing the binary image data contained in each plane of the *bitstack*.

referred to as bitplanes (Figure 4.3). The total number of bitplanes is equal to the number of binary digits used in the quantization process together with one additional plane, the *sign plane*, which stores the sign (+ or -) for each pixel. Signed images containing negative pixel values cannot be displayed because it is not possible to have negative light intensities. However, if the original image contained only real (+ve) pixels, it can be recovered by employing a weighted summation of the individual bitplane images, where the weighting factors are increasing integral powers of 2.

This method of storing digital images is important in examining and eliminating the effects of noise contamination in image signals and in determining the appropriate level of quantization which can be applied to an image. In general, the higher the bitplane number the more random is the data contained in the plane and the less significant is its contribution to the overall image structure, as shown in Figure 4.3.

4.3 DIGITAL IMAGE PROCESSING

Digital image processing is the numerical manipulation of digital images. There are two basic types of processing operation: point processing; and area (or neighbourhood) processing operations (Figure 4.4).

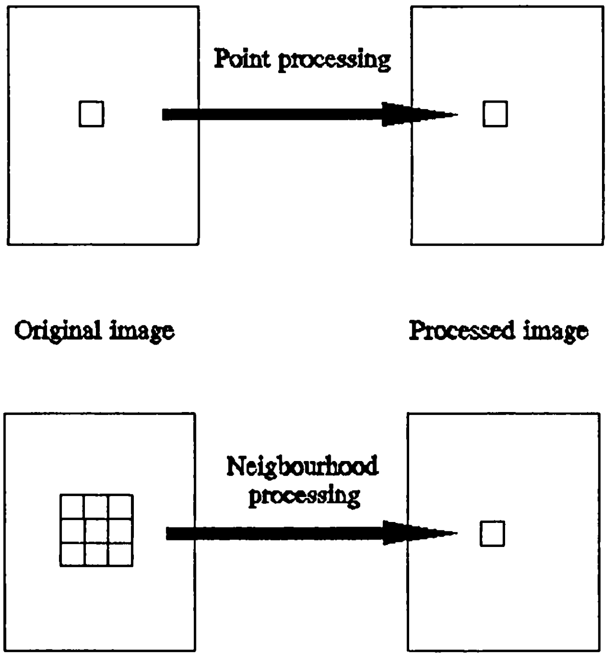


Figure 4.4 Diagrammatic representation of the transformations involved in the two basic types of image processing.

4.3.1 Point processing operations

Point processing operations consist of transforming each original pixel intensity value into a new value in the output image. The transformation depends on two factors: (i) the rule or formula to be applied; and (ii) the intensity value or grey level of the pixel in the original image. Point processing operations include numerous techniques for altering the statistical distribution of grey-levels in an image. This type of transformation is often described in terms of the first-order probability density function (PDF) or *image histogram* which plots the distribution of grey-levels in terms of the number of pixels having each grey-level. The PDF's of most natural images are approximately Gaussian with an extended tail towards the higher grey-levels (Figure 4.5).

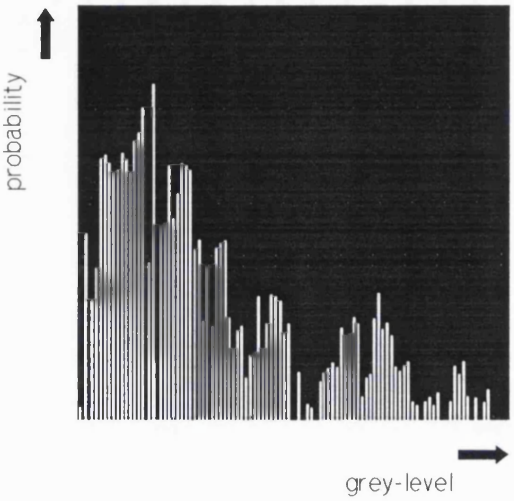


Figure 4.5 Probability density function (PDF) or *image histogram* of a natural image.

4.3.2 Neighbourhood processing operations

Neighbourhood processing operations, like point processing operations, involve a transformation of the intensity values of each original pixel into a new value in the output image. However, in this case the transformation depends not only on the grey-level of the pixel in the original image, but also on the grey levels of neighbouring pixels in the original image. This type of processing operation is therefore a context-dependent operation that alters the intensity of a pixel according to its relationship to the other pixels in the immediate vicinity. If this influence can be expressed as a weighted sum of the pixels in the neighbourhood of the pixel being processed this operation is then referred to as *linear spatial filtering* and is accomplished by the mathematical procedure known as *convolution.*

Linear spatial filtering is a technique of fundamental importance in image processing. The term "filtering" is used in analogy to electronic filters employed to select or discard radio signals of particular wavelength or frequency (for example, the tuner in a television set or radio receiver, which can be used to select a particular channel). However, in the case of an image the filter acts to select or reject particular bands of what are referred to as *spatial frequencies.*

To understand the nature of the sophisticated processing operations which may be carried out using linear spatial filtering it is necessary first to understand the way in which signals, including images may be precisely described and analyzed.

4.4 SIGNAL DESCRIPTION AND ANALYSIS

Before a signal can be analyzed or processed adequately it must be specified accurately at every point in space or time. This will typically involve approximating the signal by a mathematical expression. There are three basic methods of mathematically approximating the form of a signal: by a polynomial expansion; by a Taylor series or using a set of orthogonal functions (Lynn, 1982).

Approximating a signal by a polynomial expansion will only allow an accurate representation at (n+1) arbitrary points, where n is the order of the polynomial. The accuracy of fit can generally be improved by increasing the number of terms in the polynomial. To provide an accurate fit over the entire signal will usually require a polynomial of very high order making any analysis extremely difficult.

While a polynomial will only match the signal at several arbitrary points the second method of approximation, using a Taylor series, provides an accurate representation over a small region of the signal. As with the polynomial approximation, to achieve accuracy over a larger region of the signal a considerable number of terms are required. Consequently, approximation of a signal by a polynomial or Taylor series is only useful where it is desired to accurately represent a limited portion of the signal.

Often the most useful method of approximating a signal is by a set of so-called orthogonal functions. Simply defined, the term "orthogonal" as applied to a set of functions, indicates that each member of the set provides additional independent information about the form of the signal, such that as more members of the set are added together the approximation becomes ever more accurate. Consequently, the use of a set of orthogonal functions allows an accurate description of an entire signal.

The set of orthogonal functions most commonly used for signal description is the Fourier series or Fourier transform. In a Fourier representation a signal is broken down into a set of periodic sinusoidal and cosinusoidal functions (Figure 4.6). A signal which is characterised in this way is said to have been decomposed or transformed into its Fourier components. For a one-dimensional signal, this transformation of the signal is accomplished by applying the infinite integral of the form:

$$F(\xi_x) = \int_{-\infty}^{+\infty} f(x)\, e^{-j2\pi\xi_x x} dx \qquad (4.1)$$

Where, $F(\xi_x)$ is the one-dimensional Fourier transform of the signal, $f(x)$;

$$e^{-j2\pi\xi_x x} = \cos(2\pi\xi_x x) - j\sin(2\pi\xi_x x)$$

and j is an "imaginary number" equal to the $\sqrt{-1}$ (Papoulis, 1984). Because $F(\xi_x)$ consists of spatial frequency components (spatial sine and cosine waves), it is referred to as the *spatial frequency spectrum* or simply as the *spectrum* of the spatial function $f(x)$.

In two-dimensions, where the spatial function $f(x,y)$ is now an image, Equation (4.1) is written as

$$F(\xi_x,\xi_y) = \iint_{-\infty}^{+\infty} f(x,y)\, e^{-j2\pi(\xi_x x + \xi_y y)}\, dxdy \qquad \text{(Eq. 4.2)}$$

Where $F(\xi_x,\xi_y)$ is the spatial frequency spectrum of the image $f(x,y)$.

4.4.1 **Fourier and spatial domains of an image**

Conventional images are said to be represented in the spatial domain. That is, in the normal (x,y) coordinates of the image. Following Fourier transformation the signal is said to be represented in the Fourier or spatial frequency domain. This is another

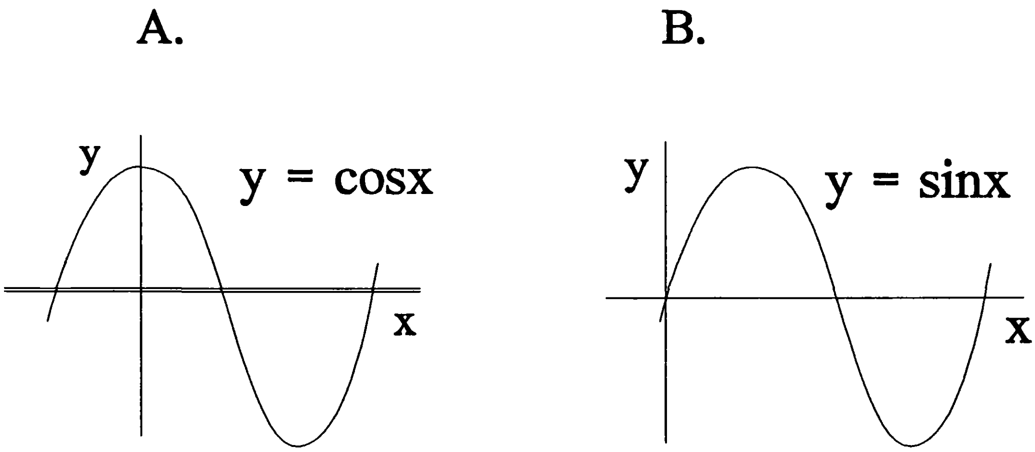


Figure 4.6 Examples of the (continuous) cosine and sine functions which form the basis of the Fourier transformation. A) Cosine function. B) Sine function.

coordinate space defined by the Fourier transform in which the coordinate axes are the frequency variables, usually written as ξ_x and ξ_y, which are the equivalents of x and y. The units of spatial frequency are cycles per unit length. To return to the spatial (image) domain the inverse Fourier transform is applied to the spatial frequency domain representation, thus:

$$f(x,y) = \iint_{-\infty}^{+\infty} F(\xi_x,\xi_y)e^{+j2\pi(\xi_x x + \xi_y y)}\, d\xi_x, d\xi_y \qquad \text{(Eq. 4.3)}$$

In general, the higher frequency components of the spatial frequency spectrum contribute to the smaller details and sharper features of an image, such as sharp edges, and the lower frequency components contribute to the large areas of shading and overall contrast of an image. It is important to note that exactly the same information is contained in the original signal as in its Fourier transform, but coded in a different way.

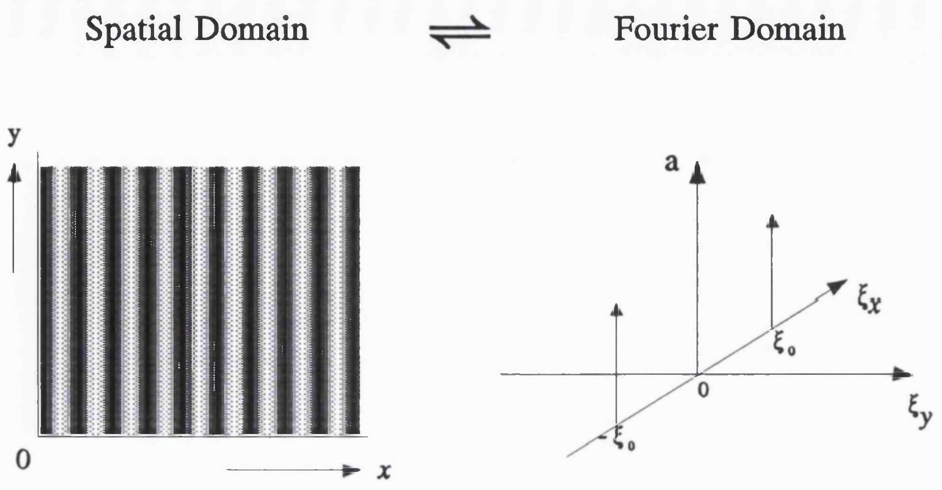


Figure 4.7 Spatial and Fourier domain representations of a two-dimensional sinusoidal function. The third axis in the Fourier domain is the amplitude of the function indicated by a.

4.4.2 **Fourier representation of discrete images**

The Fourier representation of a discrete or sampled images is different to that of a continuous signal and is referred to as its Fourier Series, $X(\omega_1,\omega_2)$ which is defined as (Jain, 1989)

$$X(\omega_1,\omega_2) \triangleq \sum\sum_{m,n=-\infty}^{+\infty} x(m,n)\, e^{-j2\pi(m\omega_1 + n\omega_2)} \qquad \text{(Eq. 4.4)}$$

Where *x(m,n)* is the sampled image in the spatial domain; and m and n are the coordinates of the sampled image corresponding to the x and y coordinates of continuous space; ω_1 and ω_2 are the discrete spatial frequency variables which are the equivalents of the discrete spatial coordinates m and n.

4.4.3 Fourier domain and spatial domain analysis of images

There are several reasons for the choice of the Fourier set of orthogonal functions the most important of which, for the present purpose, is that they are very susceptible to mathematical analysis. Thus, once the image is represented in the frequency domain it can be operated on, or *processed*, to produce the desired outcome using only basic mathematical manipulation - usually consisting of simple multiplication. However, to calculate the Fourier spectrum for even quite small images requires more than a hundred billion mathematical operations! This enormous number of calculations made the use of the Fourier transform impractical until relatively recent times when the computer based method known as the Fast Fourier Transform (FFT) algorithm was devised (Cooley and Tukey, 1965). The FFT simplifies the calculation of the Fourier transform by a factor of about ten thousand, reducing the calculation to one-hundred million mathematical operations. This is still a large computational effort.

However, many of the more complicated processing procedures are *linear* operations which only require multiplication of functions in the Fourier domain. Fortunately, multiplication in the Fourier domain has an equivalent mathematical operation in the spatial domain, known as *convolution*. The existence of spatial domain convolution allows signals to be processed in terms of their spatial frequency content but without the need for explicit Fourier transformation.

The choice between processing in the Fourier or spatial domains depends on several factors but is usually decided by the size of the functions to be processed. Regardless of which domain is chosen, the processing of images requires the complicated manipulation of large quantities of data. Because of this, it is necessary to employ an automated device or *linear filter* to process the data. The theory underlying the use of linear filters, known as *linear systems theory*, forms a central part of the processing of information contained in signals and is discussed in the following section.

4.5 TWO-DIMENSIONAL LINEAR SYSTEMS THEORY

Closely linked to the use of the Fourier representation of a signal is the processing of signals by devices known as linear systems. In the context of signal analysis and processing a system may also be defined as any device (physical or otherwise) that performs an operation on a signal (modified from Proakis and Manolakis, 1988). In the present context it is therefore a signal processing device.

Conceptually the simplest way to consider a system is as an input-output device, which supplies at its output a modified version of the input (Lynn, 1982).

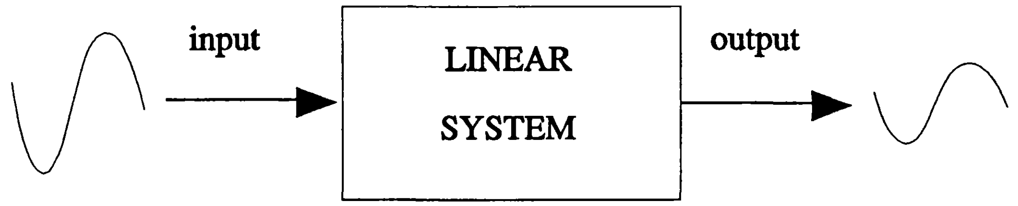


Figure 4.8 "Black-box" representation of a linear system.

In general, systems are classified as either linear or non-linear based upon the general mathematical characteristics of the type of operation that they perform on the signal. The most important class of system for signal processing is the "discrete linear" system. The theory of discrete linear systems has been developed extensively in connection with electrical circuit theory, automatic control theory and the theory of mechanical vibration, as well as signal theory; and forms the basis for the signal processing operations used in this study.

A linear system may be represented schematically, as shown in Figure 4.8, by a "black box" that receives an input stimulus. The system then produces an output response; the relationship of the response to the input stimulus is the subject of Linear systems theory. In general the relationship between input and output of a linear system is described by a "linear differential equation with constant coefficients" (it should be noted that this does not imply a straight-line or "linear" relationship between input and output of the system). More simply; if the input consists of two signals and the output consists of the sum of the outputs produced by each signal separately, the system is linear. This seemingly trivial property, known as linear superposition, has the most profound implications for signal processing. It means in particular that it is possible to evaluate separately the response to each component of a signal (Lynn, 1982).

For the present purpose there are two other important features of a linear system. The first is given by its so-called *eigenfunctions*. An eigenfunction of a system is an input function which passes through the system unchanged except for a possible alteration in its amplitude. That is, the system has no effect on the function except perhaps to alter its "strength". The eigenfunctions of a linear system are complex exponentials which include the Fourier transform and Fourier series described above.

The final feature of importance for linear systems with two-dimensional input signals, is known as *shift-invariance*. This means that a spatial shift in the input produces a corresponding shift in the output but no other change.

4.5.1 **Frequency response and impulse response**

The effect of a linear shift-invariant system on a given input signal may be described in two basic ways: by its frequency response; or by its impulse response. The frequency response is the description of the operation of the system in terms of spatial frequencies in the Fourier domain, while the impulse response is the spatial domain description of the response of the system to a infinitely powerful point stimulus known as a delta function. The impulse response and the frequency response are a Fourier transform pair, that is, one is the transform of the other. Consequently, they are equivalent descriptions of the operation of a linear system shift invariant system.

If the impulse response or the frequency response and input to a linear system are known the output can be calculated in either the Fourier or spatial domains. In the Fourier domain the frequency response and the fourier transform of the input image are multiplied together to obtain the Fourier transform of the output. In the spatial domain the output of a linear system can be found by a convolution operation between the input signal and the impulse response of the system.

4.5.2 Spatial domain convolution.

Convolution between a sampled image and an impulse response is most easily described by the use of a moving "window" - the impulse response array - passing over the image to be processed (Figure 4.9). The impulse response array is a two-dimensional grid consisting of an odd number of cells. Each cell contains a weighting factor which is multiplied with the pixel over which it lies. Thus, for a 3 x 3 window, as shown in Figure 4.9 each of the nine cells will be multiplied with the corresponding pixels producing nine numbers at any given position of the window. For each position of the window these numbers are added together. This new value is then placed at the position of the central pixel below the window in the *output* image. The window is then moved by one pixel and the procedure repeated until the window has occupied all possible positions over the input image. Consequently, each pixel in the output image is the result of a weighted summation of pixels in the neighbourhood of the corresponding pixel in the input image.

It is usually assumed that the impulse response array has a "finite region of

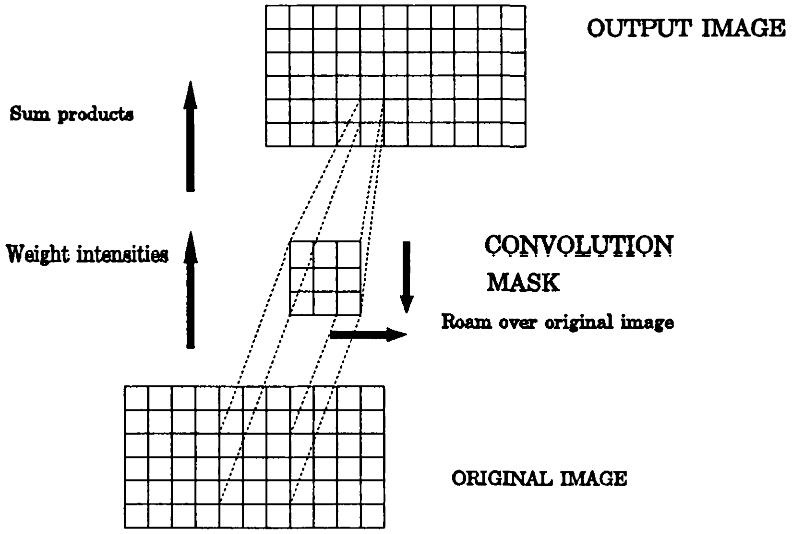


Figure 4.9 Diagrammatic representation of the convolution operation for two-dimensional images.

support". This means that a pixel in the output image is only influenced by the input image pixels over a finite area or neighbourhood of the corresponding output pixel. When the impulse response array is located on the border of the input image, the computation of the convolution does not involve all the elements of the impulse response array. This will lead to a band of pixels with spurious values at the border of the output image. This problem is known as the *border effect* or *border problem.* There are several ways of dealing with this problem. First, the convolution can be calculated as if all the missing pixels are of zero value. This is known as *centred, zero padded convolution.* Alternatively, the missing pixels can be regarded as mirror images of the pixels on the opposing border of the input image. Thus, the image is processed as if the top and bottom edges; and right and left sides, are in contact. This is known as *centred, reflected boundary convolution.*

Both, centred, zero padded convolution and reflected boundary convolution lead to visually acceptable images but for quantitative work the border pixels in the output image are meaningless. Where the output image is to be used for quantitative analysis rather than visual display, the convolution operation is only computed when the impulse response array is contained fully within the confines of the input image. Thus the output image will be smaller than the input image by the height and width of the impulse response array. This form of convolution is referred to as *centred, zero boundary convolution* and is the form used in the present study.

4.5.5. Cascade decomposition of two-dimensional impulse response arrays

An important property of finite area convolutions is that several linear systems coupled together in series produce the same overall effect as a single linear system whose impulse response is a convolution between all the individual impulse responses. Thus, a convolution between a large or complicated impulse response array can be decomposed into sequence of operations involving smaller impulse response arrays (Abramatic and Faugeras, 1982). This type of sequential operation is known as a *cascade* or *cascaded decomposition* and is very important in designing and implementing systems which large or complicated processing operations.

CHAPTER
5

THEORETICAL DEVELOPMENT OF THE DETECTION SYSTEM

5.1 INTRODUCTION

The accurate registration of radiographic images of living bone, separated by long periods of time, requires the detection of the presence and precise location of common reference points within the structure of the bone. To perform this task with any degree of certainty requires an objective comparison of the morphological information contained in the images, based on a complete mathematical description of the image data. To achieve this, the image data must first be defined by a precise mathematical function. As described in chapter 4, this is accomplished by decomposing the data into a series of orthogonal functions, the most important of which are the functions based on sinusoidal or complex exponential components - the Fourier and related transforms. Once the data have been decomposed or "transformed" in this way, they can be analysed or "processed" to extract the appropriate information - in this case, the presence and location of common structures in the bone.

The enormous complexity of such an analysis and the immense quantity of data which must be analysed, for even the smallest images, necessitates the use of an automated device or system to process the data. Such a device is known as a "filter" or "signal processor", or more simply as a "processor".

There are two distinct ways in which a processor can perform its intended task.

In the first method, the processing is carried out in the "Fourier (frequency) domain" by first decomposing the data into a series of Fourier components as described above. The transformed data are then processed using algebraic operations (multiplication and division). Finally, the processed data are converted back to the "spatial" or "image domain" (that is, the normal x,y coordinates of the original image) using the mathematical mechanism of inverse Fourier transformation. In the second method, the processing operations are performed directly in the image domain. This latter form of processing is conducted using the mathematical procedure known as "convolution" which does not require the data to be decomposed into Fourier components.

The operations necessary to process two-dimensional data can usually be specified mathematically. Because of this, and because of the enormous number of operations required, a processor of this type is most conveniently realised as a sequence of computations implemented on a digital computer.

5.2 MODEL-BASED APPROACH

The development of a processor can be accomplished in two basic ways: using a theoretical model-based, parametric method; or using a non-parametric method.

The principal advantage of model-based methods is that for physical problems well matched to the model, they generally provide a means for accurate analysis. The basic concept underlying the model-based approach is that by utilising more information about the physical processes governing the problem the more accurate will be the results. The model provides the means for formalising this information.

The model-based approach to the detection problem can be thought of as a procedure consisting of three intimately related parts (Van Trees, 1968; Candy, 1988):

(i) Model;
(ii) Criterion Function;
(iii) Processor.

The model represents the information formalising the *a priori* knowledge about the problem and indicates the quantitative information required by the processor to solve the problem. The criterion function is the mathematical rule by which the performance of the detection task may be judged. The processor is the means by which the information in the model is used to maximise (or minimise) the criterion function and thereby solve the problem. Thus, the processor represents the interaction between the model and the appropriate criterion function.

5.2.1 Formulation of the model

In order to develop a processing method using the model-based† approach, it is first necessary to define the model which will be used in the analysis.

The model must possess two fundamental properties (McGarty, 1974; Meyer, 1984). First, it must provide a means for analysis by possessing a suitable structure which is most easily met by selecting a model for which analytical solutions already exist or are known to be feasible. Secondly, it must represent a mathematical description of the physical problem of interest which can be met by identifying the essential features of the problem and choosing a tractable model which possesses these features.

The problem currently being addressed is to detect the possible presence and locations of unchanging reference sites from the x-ray attenuation patterns observed in two or more radiographs of the same bone recorded at different times. The complexity of the problem arises because of the difficulty of distinguishing between the common patterns to be detected and other similar, but unrelated, patterns which exist throughout the bone. This difficulty is compounded because the patterns to be detected are corrupted by the superimposition of other unknown and unpredictable patterns which arise from: (i) changes in the overlying bony trabeculae; and (ii) the general effects of image noise. The essential features of this problem are therefore:

(i) the common patterns to be detected;
(ii) the presence of other similar but unrelated patterns;
(iii) the unknown corrupting patterns and general noise.

---

† The term "model-based" is used here in its most general sense to mean a method based on the use of a suitable model, rather than the more restrictive sense of the "Kalman state-space methods" more common in modern signal processing usage.

Problems involving the same essential features have been extensively studied in the branch of mathematical physics known as *signal detection theory* which forms the basis of a more general field of theory referred to as *statistical communication theory* (Middleton, 1960; Sakrison, 1968). The essential features of the problem can be framed in terms relevant to statistical communications theory, thus: if the common patterns are regarded as the "signal to be detected", the similar but unrelated patterns are regarded as other "signals" and the corrupting patterns as "noise", these features can be viewed as those of a (first level) problem of classical signal detection theory - *the detection of a signal in noise* (Van Trees, 1968). The archetypal model for the problem of detecting a signal in noise is given by the equation:

$$\begin{array}{ccccc} r(x,y) & = & s_i(x,y) & + & n(x,y), \\ \\ \textit{received signal} & = & \textit{transmitted signal} & + & \textit{noise} \end{array} \qquad \text{(Eq. 5.1)}$$

Each of the $s_i(x,y)$ is a known function representing a particular possible "transmitted signal" one of which is the signal to be detected (for example, in the present case each $s_i(x,y)$ represents a pattern in the earlier radiograph); n(x,y) is the noise which in this case represents an unknown random pattern corrupting all patterns in the later radiograph; and r(x,y) is the total received signal consisting of transmitted signal and noise (in the present context this would be the bone structure represented in the later radiograph).

This general model forms the basis for two types of signal detection problem referred to as the "binary decision" and "*M*-ary decision" problems (Davenport and Root 1958). The binary decision problem is where the received signal is either signal-and-noise or noise alone; the *M*-ary decision problem is where the received signal is one of *M* possible signals each of which is contaminated with noise. The problem is to decide which particular signal was transmitted. The problem here is of the *M*-ary type.

5.2.2 Model structure for a linear processor

Having determined the general form of the theoretical model of the problem the detailed mathematical formulation, including any important assumptions, depends heavily upon the structure or class of the proposed processor (for example, a linear or nonlinear system). The restriction of the proposed processor to a particular class of systems is referred to as the *structured orientated approach* to the development of a processor (Van Trees, 1968) and is usually prompted by practical as well as theoretical considerations.

The restriction of the processor to the class of linear, spatially-invariant systems provides two important practical advantages: Firstly, the extensive tools of linear system theory become available for the solution of the problem. Secondly, in the practical application of the processor, the information formalised within the model must be quantified in some way. While the signal component of the model is assumed to be unchanging and known, the disturbances which constitute the noise component of the model are generally unpredictable, random phenomena. Consequently, there are obvious practical difficulties in adequately quantifying these complex statistical processes. However, where the processor is constrained to be a linear system, only a *partial characterisation* of the noise is required to arrive at a solution. This is because a linear system can only make use of certain limited information about random phenomena.

The partial characterisation which is most useful for a linear system is the *second-order characterisation* (Van Trees, 1968). In this representation, the noise is specified by its second-order statistics, namely, its *autocorrelation function (ACF).*

The restriction of the proposed processor to the class of linear spatially-invariant systems requires incorporation of five general assumptions into the model. These assumptions are:

(i). The forms of the signals are completely known.

(ii). The noise is a sample drawn from a random process and is statistically stationary, that is, the statistics of the noise do not vary with position in the image.

(iii). The second-order statistics (ACF) of the noise is known.

(iv). The noise is combined additively with the signal.

(v). The signal and noise are statistically independent.

5.2.2.1 *General assumptions concerning the signal*

As the signal to be detected represents those parts of the bone which are unchanging, its shape or "waveform" must be known (Helstrom, 1960). In the context of a digital image, the waveform is the intensity or transmission amplitude of each constituent pixel. The present problem limits the sources of information on the form of the signal, to the available clinical radiographs. This presents difficulties because, in practice, the image will contain unwanted disturbances in addition to the signal. These disturbances arise from many sources, principally: photographic film grain, intensifying screen grain and "mottle", projection variations, electronic-line noise, and sampling (digitisation and quantization) noise. To ensure as far as possible that the signals are known it is essential that the image is "noise cleaned" before attempts are made to detect the signals.

In addition, a more fundamental difficulty is presented by the nature of the signal itself. As the x-ray beam passes through the tissues it is attenuated by interaction with the matter along the beam path. Because of the differing radiopacity of the tissues in the path of the beam, a two-dimensional pattern of x-ray attenuation, g(x,y), is produced across the beam profile. The primary sensor (the intensifying screen(s)), makes a two-dimensional transformation of g(x,y) and the recording process, the film, makes a further transformation. These transformations are highly non-linear (as is the x-ray attenuation in the bone). Because of this and because of variations both in the spectrum and intensity of the x-ray beam it is not usually possible to recover g(x,y). It becomes necessary instead to work with the available signal in the form of the transmission intensity of the radiograph. However, while this transmission intensity pattern is related to g(x,y) this "signal" cannot be strictly considered as the "true" x-ray attenuation pattern of the tissue under study.

5.2.2.2 *General assumptions concerning the noise*

The noise is defined here as the difference between correctly registered images recorded at different times. To distinguish it from all other sources of noise in the sensing, recording and processing of the images it will be referred to as the "temporal noise".

Unlike the signal, which is assumed to be unchanging, the temporal noise is considered to be a sample from a random process which is statistically stationary. That is, the temporal noise is random and homogeneous, in much the same way as sand or pebbles on a beach (Figure 5.1). More precisely, the mean and covariance of the random process do not vary with spatial position in the image. It is therefore *covariance* and *correlation stationary*. These properties are generally referred to as "statistical stationarity".

Figure 5.1 Statistically stationary and non-stationary random images. The image on the left shows the homogeneous distribution which defines statistical stationarity. The image on the right is non-stationary.

The second-order statistics (the autocorrelation function) of the temporal noise are assumed to be known. This presents some fundamental difficulties because of the practical definition of the temporal noise. That is, because the temporal noise is defined to be the difference between correctly registered images, large samples of registered image data are required to estimate these statistical properties. In general, these data will not be available. Therefore, to apply the processor to any given situation requires that the autocorrelation function or mean and covariance are estimated from a theoretical model. A three-dimensional mesh plot of the ACF of an image is shown in Figure 5.2.

Together with the three general assumptions an additional assumption is made, which, although not strictly necessary, nonetheless facilitates analysis. That is, the (first-order) probability density function of the temporal noise is Gaussian with a mean value of zero.

5.2.2.3 *Assumptions concerning the association between signal and noise*

The temporal noise is assumed to be additive and statistically independent of the signal. Noise which depends upon the level of the signal or is combined with the signal in a multiplicative way is not considered by this model.

For the problem considered here there are four sources of "noise" which may violate this assumption: Quantum noise from the original radiographic exposure; film-grain and intensifying screen granularity; and geometric distortion of the image due to differences in projection.

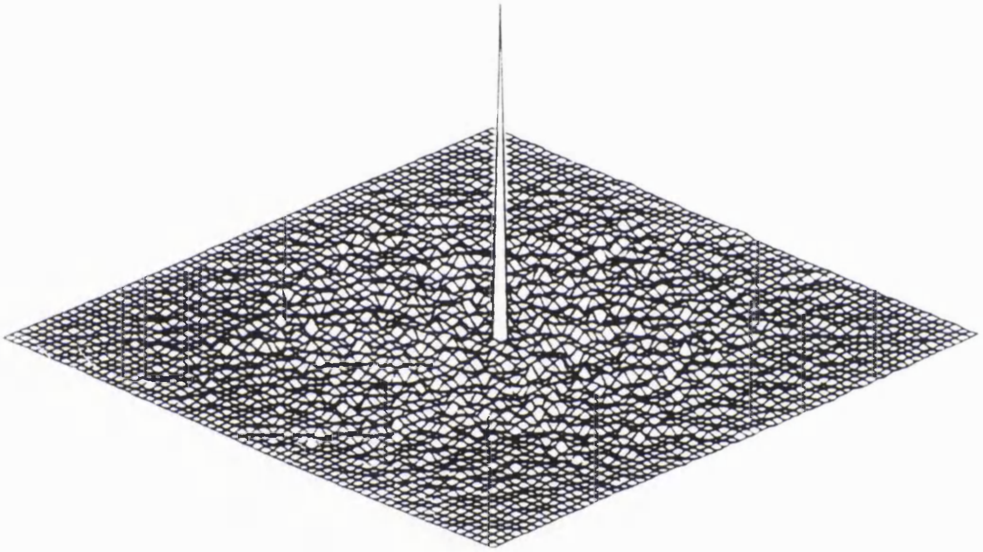


Figure 5.2 Three-dimensional mesh plot portraying the Autocorrelation Function (ACF) of a random image.

The Quantum noise and film-grain and screen noise can only be dealt with entirely satisfactorily by defining a more complicated model to accommodate signal-dependent noise. While a more elaborate model of the temporal noise is likely to be more realistic, the incorporation of a signal-dependent noise term leads to considerable complications in the mathematical analysis of the model (Jain, 1989; Pratt, 1991).

To avoid these problems and still comply with the assumptions of additivity and statistical independence requires some preliminary amendments to the problem. First, the temporal noise is constrained to be additive by defining it *a priori* as the arithmetic difference between the earlier and later images; and secondly, statistical independence of signal and temporal noise is achieved by separating out the strongly signal-dependent components, and dealing with them individually. In this case the most strongly signal-dependent elements of the temporal noise are the quantum noise and geometric distortion.

Quantum noise arises from the statistical fluctuations in the numbers of x-ray quanta reaching the intensifying screens. The fluctuations are visible because of the low level of x-radiation used to produce a clinical radiograph. To reduce the effect of quantum noise a "noise cleaning" procedure can be applied to the digital image acquired from the radiograph. This consists of smoothing the fluctuations by convolving the image with a local averaging mask or median filter (Pratt, 1991).

The second source of signal-dependent noise is the geometric distortion which arises from the differences in projection geometry between the recording of each radiograph. The assumption is made that the differences in projection are small and this contribution can be approximated, therefore, by an affine distortion of the signal coordinates (Mostavi and Smith, 1978b). This contribution is therefore highly signal-dependent and increases with the image size. As explained above, such signal-dependent noise is difficult to handle analytically and is excluded by considering only small subsections of the imagery where the distortion between images is small enough to be ignored. The basis for determining what magnitude of distortion can be neglected is its effect on the registration error variance. Mostavi and Smith (1978b) have investigated the effects of affine distortion on the registration error variance for three dimensional scenes with small depth relative to sensor distance. They found the optimum subimage size to depend on the angular distortion and the scene correlation length. For small isometric distortion ($\leq 5°$) the optimum subimage shape and size to give minimum local registration error variance is a square between 15 and 30 pixels on a side.

5.2.3 Critical comments on the model structure

The archetypal model derived from statistical communication theory yields a process which is much simpler than the actual situation it attempts to describe. For instance, because of the large dynamic range of clinical x-ray images, the signal dependent nature of the screen and film-grain noise is likely to be relatively more important than in other photographic recording systems to which this model may be applied. Consequently the temporal noise will almost certainly contain elements which are more appropriately modelled as multiplicative noise rather than entirely additive. Furthermore, the statistical stationarity of the temporal noise may be questionable especially where large changes have occurred between the images.

Another aspect which has been entirely neglected is the possibility that a different class of model might approximate the situation rather better than the simple additive noise model. For example, a model in which both earlier and later images are contaminated with noise or in which the effects of geometric distortion are specifically accommodated.

Accordingly, the theoretical model defined here may only have limited validity for the present problem. Nevertheless, the model appears to provide a reasonably accurate representation of the physical situation. Although the addition of further features to the model would make it more realistic, the price to be paid for such an improvement in the model is likely to be excessive in terms of the mathematical complexity.

5.3 THE CRITERION FUNCTION AND DETECTION PROCESSOR

The model derived in the previous section represents the information formalising the *a priori* knowledge about the problem and indicates the quantitative information required by the processor to solve the problem. This section describes the remaining aspects of the model-based approach to developing a detection processor: the criterion function and the processing algorithm.

5.3.1 The criterion function

Criterion function is a measure of how well the processor performs its intended task. It is a mathematical rule by which the performance of the detection task may be judged. The criterion function can take many forms but is usually selected on the basis of information which is meaningful about the problem. However, the precise form of the criterion function is heavily influenced by the ease with which a processor may be developed. Criterion functions applicable to detection and estimation problems may be classified as either deterministic or probabilistic. Because one of the inputs (temporal noise) is a sample function from a random process it is appropriate to select a probabilistic criterion function.

As the processor is constrained to be a linear spatially-invariant system, a criterion function is selected from those appropriate to this class of system. There are four well known criterion functions relevant to linear systems and which lead to workable analyses:

Mean-mean square error criterion;
Signal-to-noise criterion;
Likelihood ratio criterion;
a posteriori probability criterion;

The appropriate criterion function depends on the precise formulation of the problem embodied in the general model. Referring again to the problem at hand: if the set of subimages, s_m, from the earlier radiograph, are viewed as the set of signals representing the possible stable sites or structural features in the bone, then, for a given subimage, r, from the later radiograph, we wish to determine which signal, s_m, m=1,2, ...,M has the maximum probability of representing the same portion of bone. Stated more formally the processor is directed to select the signal/subimage, s_m, which maximises the *conditional probability*, $p(r \mid s_m)$. Stating the problem in this way requires us to employ the likelihood ratio as the criterion function.

The theoretical form of the linear processor which employs the likelihood ratio and therefore provides an output that maximises the conditional probability, $p(r \mid s_m)$, was originally derived in the context of pulsed radar target detection (North, 1963) and is known as the *conjugate* or *matched filter*. The terms "conjugate" and "matched" arise because the filter (a "filter" is a frequency selective processor) is matched to the signal it is attempting to detect by using a mathematical form known as the "complex conjugate" (Turin, 1960).

The matched filter detection processor can be thought of as a statistical decision device where the problem is to determine which signal, s_m, m=1,2, ...,M has the maximum probability of representing the same portion of bone. This can be formulated in terms of M hypotheses:

$$
\begin{aligned}
H_1 : r(x,y) &= s_1(x,y,) + n_r(x,y) \\
H_2 : r(x,y) &= s_2(x,y,) + n_r(x,y) \\
H_3 : r(x,y) &= s_3(x,y,) + n_r(x,y) \\
\cdot \\
\cdot \\
H_m : r(x,y) &= s_m(x,y,) + n_r(x,y)
\end{aligned}
$$

In order to determine which of *M* discrete signals gave rise to the observations, it is necessary to accept the hypothesis, H_m, for which $\varphi_m\ p_m\ (r)$ (m=1, 2, ... M) is the greatest, Where $p_m(r)$ is the probability that if the *m*th subimage (signal) represents the site of stable detail, then r(x,y) will be observed in the later image. The φ_m's are constants, independent of r(x,y), which are determined solely by the criterion of the hypothesis test. For the class of linear spatially invariant systems the four criteria above lead to the derivation of the same processor. If the chosen criterion is the minimisation of the overall probability of error then φ_m is the *a priori* probability that the *mth* subimage represents the stable bony region. If the *a priori* probability is not known the φ_m are all equated to unity. In this case the test is based upon maximising the likelihood criterion and the test is called a "maximum likelihood test". A processor which performs such a test may be referred to as a "maximum likelihood processor". **Thus, for any given subimage in the first image, the maximum likelihood (ML) processor selects the subimage in the second image with the greatest probability of representing the same portion of bone. That is, it maximises the likelihood ratio and the conditional probability, $p(r \mid s_m)$.**

5.4 THE MAXIMUM LIKELIHOOD DETECTION PROCESSOR

The mathematical operations required to perform ML detection can be determined by solving the problem in either the frequency domain or in the spatial domain. The form derived in the frequency domain is purely algebraic and easier to obtain than the alternative spatial function (Cook and Bernfield, 1967). Having obtained the frequency domain function it can then be converted back to a spatial function to determine the operations which must be performed on the images to implement the processor.

The "frequency response" for a two-dimensional processor (that is, the mathematical description of its mode of operation in spatial frequency terms) is given by the so-called *matched filter equation*:

$$H(\omega_1,\omega_2) = \frac{S^*(\omega_1,\omega_2)}{|\,N_t(\omega_1,\omega_2)\,|^2} \exp[-j(\omega_1 m,\omega_2 n)] \qquad \text{(Eq. 5.3)}$$

where:

$H(\omega_1,\omega_2)$	is the frequency response of the processor;
$S^*(\omega_1,\omega_2)$	is the complex conjugate of the signal;
$\|\,N_t(\omega_1,\omega_2)\,\|^2$	is the (power) spectral density of the noise;
$\exp[-j(\omega_1 m,\omega_2 n)]$	is the exponential factor indicating the Fourier domain;
ω_1 and ω_2	are the spatial frequency coordinates corresponding to the horizontal and vertical spatial coordinates (i.e. x and y).

5.4.1 The operation of the ML processor

The operations performed by the ML processor depend on whether or not the noise contaminating the signal, is "white" (Figure 5.3). "White" or "uncorrelated" noise is defined as a random disturbance with a uniform spatial frequency spectrum. This form of noise image can be seen as the snow-like appearance on a television screen which is tuned to a channel that is not broadcasting.

Conversely, noise which does not have a uniform spatial frequency spectrum is referred to as "coloured" or "correlated" noise. The spatial frequency spectrum of coloured noise is typically characterised by a decline in power with increasing spatial frequency. In this case, the corresponding effect in the spatial domain is a statistically correlated image, that is, an image which has spatial structure.

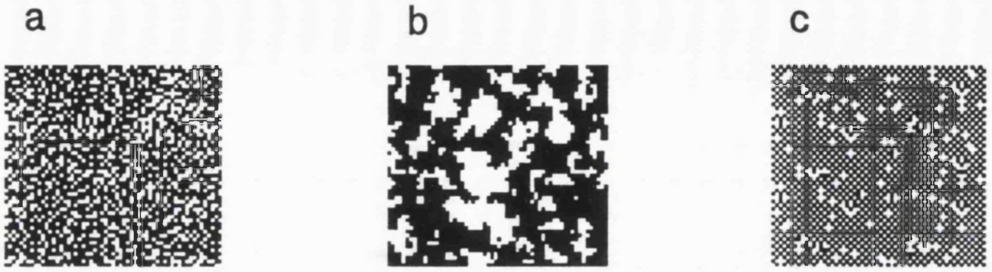


Figure 5.3 Binary images depicting random noise: (a) uncorrelated ("white") noise; (b) correlated (low frequency) noise; (c) correlated (high frequency) noise.

5.4.1.1 *The white noise case*

Where the noise is white the denominator in equation 5.3 is equated to unity and under these circumstances the operation performed by the ML processor is the mathematical process of cross-correlation of the signal to be detected with each subimage of the image being searched for a match (Cattermole, 1986). The subimage with the greatest probability of being the correct match, that is, representing the same portion of bone, is indicated by the spatial position at which the cross-correlation reaches its maximum value. A three dimensional plot of the output of a detection processor is shown is Figure 5.4 where the conditional probability is plotted on the vertical axis against the spatial position on the image where the match is being sought (indicated by the two horizontal axes, x and y).

5.4.1.2 *The correlated noise case*

Where the noise accompanying the signals is correlated the operation performed by the ML processor is rather more complicated than in the case of white noise. Under these circumstances the ML processor operates by performing preliminary processing to transform the noise into white noise and then proceeds in the same manner as for the white noise case indicated above (Kailath, 1970). The operation of the ML processor for correlated noise can be thought of as two linear processors acting in series (Figure 5.5).

The first processor transforms the correlated noise to uncorrelated (white) noise by equalising the power at all spatial frequencies. This process is known as *whitening* or

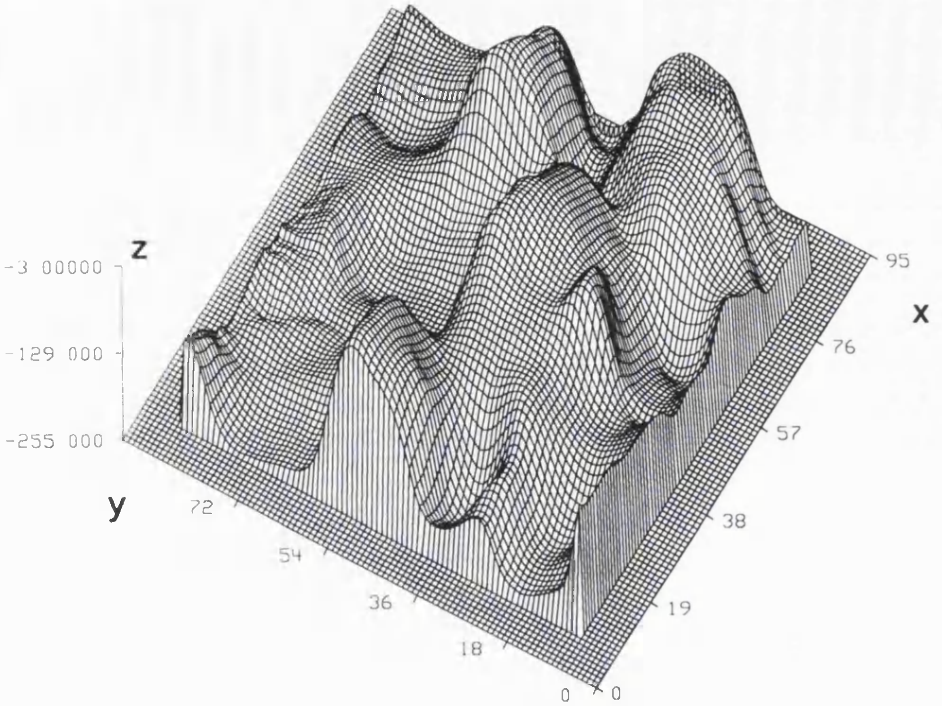


Figure 5.4 Three-dimensional mesh plot showing the output of a detection processor. The cross-correlation value is plotted on the vertical axis against the spatial position on the image where the match is being sought (indicated by the two horizontal axes, x and y).

decorrelation. The second processor then detects the signal (which is now submerged in white noise) using cross-correlation as described above. However, because the signal and noise are intimately bound together, the first processor acts on, and alters, both the noise and the signal. Accordingly, the output from the first processor consists of white noise plus a spectrally distorted signal. Consequently, the second processor must be designed to detect the spectrally distorted or "whitened" signal rather than the original signal. This is accomplished by passing both sets of image data (from earlier and later radiographs) through the first processor. As before, for any given subimage in the first radiograph, the overall processor selects the subimage in the second radiograph with the greatest probability of representing the same portion of bone. Again, this is indicated by the spatial position at which the cross-correlation, performed by the second processor, reaches its maximum value.

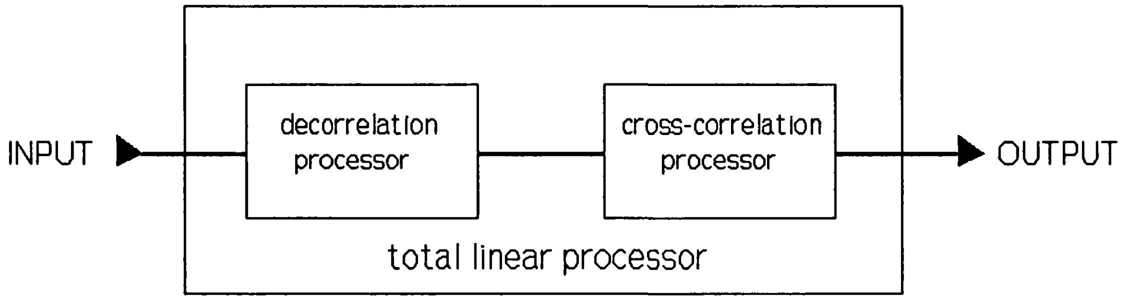


Figure 5.5 Diagrammatic representation of a two-stage detection processor for signals in correlated noise.

5.5 PROCESSOR INTERPRETATION

The ML detection processor is a device which detects common structures by recognising and quantifying similarities in the spatial frequency spectra of their two-dimensional x-ray attenuation patterns. This task is performed by cross-correlating the noisy received signal (a subimage of the later radiograph) with each transmitted signal (the possible matching sites in the earlier radiograph). As each possible matching signal is compared with the received signal the processor produces an output which is proportional to the probability that the two signals represent the same section of bone. The signal which produces the greatest output has the highest probability of being the correct match.

Where the noisy signal is described by the mathematical model of Equation 5.1,

$$r(x,y) = s(x,y) + n(x,y)$$

the cross-correlation of the signal-plus-noise, r(x,y) with the reference signal, s(x,y), is given by the equation,

$$R_{sr} = R_{ss} + R_{ns} \qquad \text{(Eq.5.4)}$$

where R represents the correlation between the functions denoted by the subscripts (e.g. R_{sr} is the (cross-) correlation between s(x,y) and r(x,y)). If the noise, n(x,y), is uncorrelated ("white") and independent of the signal, s(x,y), the cross-correlation between signal and noise will, on average, equal zero. Thus, the cross-correlation of signal-plus-noise, r(x,y) and signal, s(x,y), will be the same as the autocorrelation of the reference signal with itself (that is, $R_{sr} = R_{ss}$). An important property of autocorrelation functions is that they have a maximum value at zero-displacement, that is, when the two copies of the signal exactly overlie one another. Since $R_{sr} = R_{ss}$ the cross-correlation, R_{sr}, will also contain a peak at the point where the noisy signal and reference signal exactly overlie. This behaviour permits the detection of the presence and location of the noisy signal.

However, if the noise is internally correlated there are likely to be parts of the spectrum which are shared or held in common by both noise and signal, even though the signal and noise are effectively independent. Thus, the cross-correlation of signal and noise will no longer be zero and there can be no guarantee that the peak of the cross-correlation will occur at the point where the reference signal and noisy signal overlie one another. This is especially likely to be a problem where the signal and noise have similar spectra. For the processor to operate correctly, it is essential that the signal is only contaminated by statistically independent white noise. This can be achieved by passing the signal-plus-noise through a whitening filter or decorrelator prior to performing the cross-correlation.

Unfortunately, even if the signal is accompanied by white noise it may be difficult to locate the peak of the cross-correlation where the shape of the cross-correlation is broad and flat rather than sharply peaked. A large sharp peak only occurs where the signal has a spatial structure similar to that of white noise. So, in addition to whitening the noise it is helpful to whiten the signal as well. Where the noise and signal have similar spatial frequency spectra this will occur automatically as the noisy signal is passed through the decorrelator. Accordingly, the decorrelator performs two tasks: it converts the correlated noise to uncorrelated (white) noise; and it alters the structure of the signal allowing it to be located more easily.

This theoretical interpretation of the action of the processor is rather simplistic. In practice, the cross- and autocorrelation functions employed by the processor, must be calculated from finite observations which makes the detection of the correlation peak uncertain, particularly where the signal-to-noise ratio is low. The problems of finite samples places practical limitations on the operation of any detection processor which employs cross-correlation.

5.5.1 Conclusions

The theoretical form of the detection processor has been developed using a model-based approach. This has necessitated the adoption of several simplifying assumptions, both in the basic model and in the subsequent development. The more important of these assumptions must be verified empirically and this is undertaken in chapter 8. The mode of operation of the detection processor has been defined and explained in terms relevant to signal detection theory.

CHAPTER
6

THE THEORETICAL MODEL OF THE TEMPORAL NOISE

6.1 INTRODUCTION

The development of a signal processor for the detection of common reference sites revealed in serial radiographs of growing bone has been approached from the perspective of statistical communication theory. A central feature of this approach is the characterisation of the change that has occurred between the recording of the radiographs - the temporal noise. It is this noise which is the major cause of the uncertainty in detecting the common reference sites. Although the temporal noise in a particular pair of images is unpredictable, knowledge of the statistical characteristics of the noise will increase the information available to the processor. The purpose of this chapter is to provide a theoretical model of the statistical properties of the temporal noise which can later be used to develop the so-called "whitening filter" or decorrelator.

6.1.1 The approach to characterising the temporal noise

The implementation of the detection processor to be used in a particular practical situation requires the statistical characterisation of the temporal noise. The nature of the characterisation is determined by the particular approach which is taken. Where the processor is constrained to be a linear system, the required characterisation consists of the second order statistics: that is, the spatial autocorrelation function (ACF) of the temporal noise.

Knowledge both of the form and the parameter values of the ACF are needed to develop a practical processor. This presents a fundamental problem because the temporal noise is defined as the arithmetic difference between transmission amplitudes of the earlier and later images, which requires that the true position of registration is known. If this was the case there would be no need for the processor.

One solution to this difficulty would be to estimate the temporal noise ACF empirically by examining similar image pairs whose registration is already known (Svedlow, McGillam and Anuta, 1978). For example, by subtracting an earlier radiograph from a later one using the positions of implanted markers to define the relative orientation or registration position. There are two weaknesses to this approach. First, for many of the possible applications of the processor, the large numbers of accurately registered radiographs are not likely to be available; and secondly, the ACF deduced in this way is purely descriptive with no obvious association with the physical processes underlying the temporal noise, thereby making it difficult to extrapolate the findings to other different sets of radiographs. An alternative approach has therefore been adopted.

First, a simple model is developed from a set of plausible physical properties of the temporal changes which constitute the noise. Secondly, the relevant features are identified and formalised in a mathematically tractable yet simple statistical model of the temporal noise. Finally, the form of the ACF is then deduced from the statistical model.

6.2 THEORETICAL MODEL OF THE TEMPORAL NOISE

Throughout this section the general assumptions stated in section 5.2.2 are deemed to apply. That is, that the temporal noise is the difference between small, correctly registered, subimages of the earlier and later radiographs with minimal differences in relative geometric distortion. The temporal noise is statistically stationary and independent of the signal. In addition, the radiographic images from which the temporal noise is derived are assumed to have been sampled to provide digital images which have subsequently been radiometrically corrected to give similar mean and variance to the intensity values.

The starting point of the theory is the assumption that the temporal noise is composed of three statistically independent component images. The first contribution is due to the background noise which arises from the differences in sensing, recording and converting the x-ray attenuation patterns into manipulable images. It comprises the differences between the images which are essentially independent of alterations in the tissues being imaged. This component will be referred to as the "background image". The second contribution is due to the variable superimposition of structures deep and superficial to the stable structures which constitute the common reference points. This component will be referred to as the "structural image".

The third contribution is due to changes in the bone structure. That is, the remodelling and restructuring of the bone that has occurred between recording the two radiographs. This component will be referred to as the "growth image".

The background image, structural image and growth image are assumed to combine additively to form the total temporal noise.

6.2.1 The background image

The background image occurs because of differences in the recording of the images acquired from the original x-ray attenuation patterns. Thus, even if the two patterns were identical the pixel arrays will not be the same. The background image consists of two major parts: radiographic imaging noise; and digital image acquisition noise.

Radiographic imaging noise is the noise injected into the radiographic image by the nature of the radiographic process. It arises from two principal sources: x-ray quantum noise - the statistical variations in the x-ray flux reaching the primary detector (intensifying screen); and from the random distribution of grains in the screen and photographic film.

The x-ray quantum noise results from the desire to keep the radiation exposure to the tissues as low as possible while still allowing the production of a clinically useful image. This is usually assumed to obey Poisson rather than Gaussian statistics, and is signal dependent. However, in calculating the total noise for x-ray screen/film systems it is usual to assume that the quantum noise is statistically independent of the signal, uncorrelated and additive. The justification for this is that predictions based upon these assumptions are in close agreement with experimental data (Dainty and Shaw, 1974). The total radiographic noise is dominated by the x-ray quantum noise at spatial frequencies below approximately 4 or 5 cycles/mm; above this level film grain noise predominates (Barrett and Swindel, 1981b). The contribution of the x-ray quantum noise to the background image will be the difference between two realisations or samples of this noise (one for each of the two radiographs). Given that quantum noise is statistically stationary and assuming the x-ray exposures are similar, the resulting noise will be uncorrelated, zero-mean and approximately Gaussian since the Poisson distribution approaches a Gaussian form for large samples.

The second source of radiographic imaging noise is the screen-grain and film-grain noise which originates from the random distribution of grains in the intensifying screen and photographic film. Screen-grain noise is usually negligible in comparison to film-grain noise (Barrett and Swindel, 1981b). Although film-grain noise is weakly signal dependent it is typically considered to be statistically independent of the signal (Jain, 1989). It is also an uncorrelated, statistically stationary Gaussian random process with a mean value of zero (Doi, 1969).

Digital image acquisition noise consists of two types of noise: digitization noise; and quantization noise. Digitization noise occurs because the acquisition of the digital images necessitates the averaging of small regions of transmission intensity (image density) data. This gives rise to random differences between the images because of the statistical nature of the sampling process and because the averaging is not carried out over precisely the same anatomical regions. Digitization noise is random and spatially uncorrelated. Quantization noise is generated by the discrete steps in the analogue to digital conversion of the original radiographic image (Pratt, 1991). That is, while the original radiograph has a continuous brightness range, the brightness of the digital image can assume only integer values. Where the radiographic brightness is at the decision level between two quantization bands its allocation to one or other band is essentially random. Quantization noise is therefore considered to be zero-mean, uncorrelated noise.

These sources of noise (digitization and quantization) are characterised by random statistical variations in the grey-level from pixel to pixel in each original image. They are uncorrelated, Gaussian and statistically stationary with a mean of zero. They are also independent of the signal.

To determine the statistical form of the background image requires the estimation of the difference between two images containing these forms of noise. The statistical structure of the background image can be estimated using a fundamental property of Gaussian random processes. That is, the combination of independent Gaussian random variables yields a further Gaussian distributed variable with a variance equal to the sum of the respective variances and a mean which is equal to the sum or difference of the respective means dependent on the nature of the combination (Papoulis, 1984). The background image, the difference between two samples of these noises, will be consequently of the same form as the individual component noise sources. That is, a zero-mean Gaussian random image in which the pixel values are mutually uncorrelated. This form of random image is referred to as a zero-mean white Gaussian field.

6.2.2 The structural image

The second contribution to the temporal noise is due to the variable superimposition of structures deep and superficial to the stable structures which constitute the common reference points. The structural details (usually referred to as the "trabecular pattern") revealed on radiographs of the craniofacial bones are generally created by the radiodense junction between the cancellous bone and the endosteal surface of the cortex (Ramadan and Mitchell, 1962; Pauls and Trott, 1966). Thus, while small changes in the projection geometry may cause no observable geometric distortion in any given structure because it occupies only a limited depth, the relative positions of other structures along the x-ray path will be altered because of the greater increase in disparity with increasing depth (parallax).

The form of the structural image will depend heavily on the spatial form of the trabecular pattern. Although the specific form of the trabecular pattern can be expected to vary from individual to individual and from bone to bone, the general form the radiographic trabecular pattern is a network of large sparse orthogonal elements with numerous smaller off-orthogonal and diagonal elements (Oxnard, 1982; Ishizuka, 1981; Van der Stelt and Geraets, 1987). Consequently, the radiographic trabecular pattern generally shows clearly defined linear elements with distinct and preferential directions of orientation. Unlike the components of the background image the trabecular pattern is strongly spatially correlated. That is, the intensity at two or more points on the image of a bony element will be statistically related.

The structural image - the difference between two images with slightly altered projection geometry - will be composed of the differences between the overlapping layers of the trabecular pattern. For small changes in projection, the difference between the two images will arise from a partial subtraction of each pattern from itself and a partial subtraction from other overlying patterns. Because of the multiple overlapping patterns involved in the production of the difference image, this component of the noise will appear (approximately) randomly distributed with a mean intensity of zero. For large samples this component can be expected to have a Gaussian distribution.

The multiple partial subtractions of the spatially displaced patterns will give rise to a random image in which some of the spatial correlation of original trabecular pattern will be preserved. Consequently, the structural image will be a spatially correlated zero-mean Gaussian random field.

6.2.3 The growth image

The growth image, the final contribution to the total temporal noise, is due to true changes in the bone structure - the remodelling and restructuring of the bone that has occurred between recording the two radiographs.

Remodelling and restructuring of bone must cause substantial changes in the mineral content before they become visible radiographically either visually or using sophisticated quantitative analysis (Troubach et al., 1984). Therefore, although there is a constant turnover of calcium in living bone, this occurs without appreciable structural change (Parfit, 1979). However, over prolonged periods of time the physical structure of bone can alter substantially especially at times of rapid general growth (Jerome and Gubler, 1991). In addition, there is a more rapid restructuring in certain disease states, for example, hyperparathyoidism (Frost, 1973). These changes take place at many scales from the molecular level to the macroscopic level of the whole bone. However, the character of these changes is highly dependent on the spatial resolution at which these events are observed. The levels of interest for the present problem are those visible in clinical radiographs, for which the lower limit is approximately 0.05 - 0.25 mm dependent on the type of radiographic image.

One useful source of information on the changes in the structure of bone at this range of spatial sizes is the experimental investigation of mammalian bone by intra-vital staining as revealed under low power microscopy. The work of Hoyte (1968), and Frost (1969) on growing bone using Alizarin and fluorescent tetracycline markers respectively, revealed two main features which are of interest in developing a theoretical model of the growth image. These were: (i) local spatially (and temporally) correlated deposition, resorption and stasis; and (ii) the site of change was essentially restricted to the surface of the trabeculae and subperiosteal area.

Given the structure of the trabecular pattern discussed above (section 6.2.2.) the effects of remodelling and restructuring of bone will be to alter the spatial correlation of the trabecular structure. Thus where the projection geometry is constant, the growth image is likely to have a similar form to the structural image. However, changes in projection geometry, which inevitably occur between succeeding radiographs, will cause areas of remodelling and restructuring to be superimposed on (and subtracted from) other different regions of bone. As with the structural image, the many overlapping patterns involved in the production of the difference image will lead to an (approximately) Gaussian random image with a mean intensity dependent on the ratio of calcium loss to calcium gain. For short periods of time, and in the absence of major bone pathology, the mean change in calcium content will be approximately zero (Polig and Jee, 1987). The growth image will therefore appear as a spatially correlated zero-mean Gaussian random field.

6.2.4 The relative contribution to the total temporal noise

The background image, structural image, and growth image are assumed to be statistically independent combining additively to form the total temporal noise. If the statistical properties of the images were the same, their relative contribution to the total temporal noise would be unimportant because its structure would not be altered. However, where the statistical properties of the component images are different their relative contribution will be of major importance in establishing the statistical properties of the temporal noise. Therefore, it is important to establish the relative contributions of the background image and the combined structural and growth images. The structure of the component images are summarised in Table 6.1.

In the knowledge that serial clinical radiographs provide information of a

Table 6.1 Components of the Temporal Noise.

Image Field	Underlying Process	Statistical Structure
Background Image	Quantization noise	uncorrelated Gaussian zero-mean (white noise)
" "	Digitization noise	uncorrelated Gaussian zero-mean (white noise)
" "	Variations x-ray flux (Quantum noise)	uncorrelated Gaussian zero-mean stationary. Weakly signal dependant (white noise)
" "	Film-grain noise (Screen noise)	uncorrelated (?) Gaussian zero-mean, stationary. Weakly signal dependant (white noise)
Structural Image	projection differences	spatially correlated zero-mean Gaussian random field.
Growth Image	remodelling and restructuring of bone	spatially correlated zero-mean Gaussian random field.

reasonable character for longitudinal study, it is clear that the background image will generally be of low intensity and its effect is likely to be minor in relation to the other sources of noise. In addition, the background image is assumed to be independent of the structures being imaged and is, therefore, approximately constant for any given set of sensing, recording and converting systems. This is possibly also true for the structural image because it is related to the variations in projection geometry which are assumed to be random, and therefore unlikely to be affected by the interval between radiographs[†]. Conversely, the contribution from the growth image will depend on the time interval between recording the radiographs. It is reasonable to assume that the contribution from the growth image will generally increase as the interval between the radiographs increases. Consequently, for actively growing bone which is subject to rapid and continuous restructuring, the major contributions to the total temporal noise are likely to arise from the structural and growth images. However, the relative contribution of growth and structural images is assumed to be unimportant because of their similar statistical structure.

6.3 THE STATISTICAL MODEL OF THE TEMPORAL NOISE

Having established that the characteristics of the temporal noise will be determined largely by the combined structural and growth images, the statistical structure of the temporal noise can be inferred to be a spatially correlated random image with an approximately zero-mean Gaussian distribution. However, to determine the theoretical form of the ACF it is helpful to have an specific statistical model which displays the same general characteristics as the temporal noise and which is analytically tractable in terms of: (i) determining the ACF; and (ii) developing the decorrelation filter.

6.3.1 The Choice of Statistical Model for the Temporal Noise

Although the temporal noise is primarily random, the inherent spatial correlation indicates a degree of statistical dependence between the constituent pixels. This type of discrete random image can be approximated by members of the class of mathematical models known as "discrete random spatial interaction or "random field" models (Kashyap, 1981). However, the random field model known as the *first-order Markov random field* (MRF) explicitly captures this important property of local spatial dependence and its ACF has a well defined form with few parameters. In addition, the MRF offers an analytically

---

[†] There is some, albeit circumstantial, evidence that variations in projection geometry probably increase with intervals $>$ 2 years between recording successive radiographs (McWilliam, 1983).

tractable mechanism for developing a decorrelation filter to decorrelate the noise.

6.3.2 The Markov random field model

A random field is defined as a two-dimensional grid of random variables, where each random variable can have only one of K values or levels (0,1,2,..., K-1). A Markov random field (MRF) is defined as a random field with three distinctive properties (Cross and Jain, 1983): Positivity; Markovianity; and Homogeneity. Simply stated these properties mean that the value of each pixel in an MRF is statistically dependent on the values of its neighbouring pixels; and this dependency does not vary with position in the image. A binary MRF is shown in Figure 6.1 together with an uncorrelated random field for comparison. As can be seen, the MRF image reveals spatial structure not apparent in an uncorrelated random field which is essentially uniformly distributed.

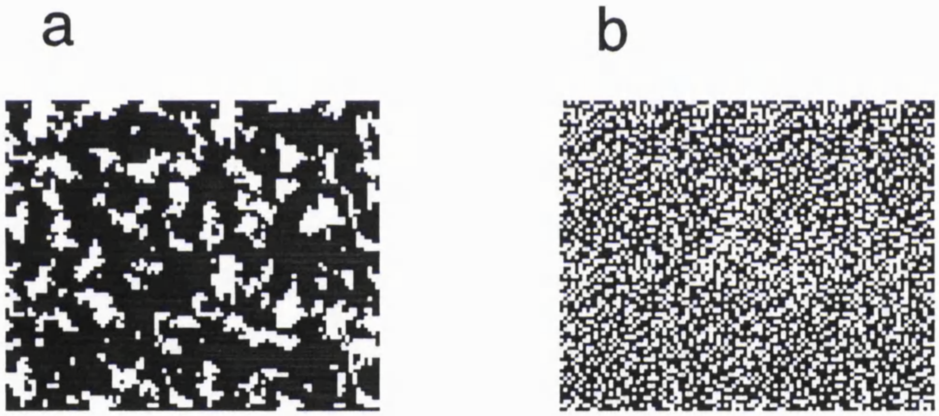


Figure 6.1 Binary random images: (a) Markov random field; (b) uncorrelated (uniformly distributed) random field.

The order of the MRF is defined by the number of neighbouring pixels considered for the conditional probability. For example, in a first-order MRF the four nearest pixels are considered and in a second-order MRF the eight nearest pixels are considered.

To ensure that the temporal noise model is applicable to a wide range of images and is practical, in the sense that it requires the estimation of only a minimum number of parameters, the first-order MRF (the simplest MRF), has been chosen to represent the temporal noise. Since the temporal noise is considered to have a zero-mean Gaussian distribution of the pixel values, the temporal noise model is therefore a *zero-mean Gaussian first-order MRF*.

6.3.3 Analysis of the form and parameters of the ACF

The statistical characteristics of the temporal noise which are of importance to the development of the processor are the form and parameters of its autocorrelation function (ACF). There are two forms of ACF applicable to first-order MRF's: separable and nonseparable. The separable form implies that the correlation between the pixels is separable into the product of horizontal and vertical correlations, while the nonseparable form does not imply this property. For a random image with a mean value of zero the ACF's are given by the following equations (Jain, 1989):

(Separable form)

$$R(m,n) = \sigma^2 \rho_1^m \rho_2^n \qquad \rho_1 < 1, \ \rho_2 < 1 \qquad \text{(Eq. 6.1)}$$

(nonseparable form)

$$R(m,n) = \sigma^2 \exp\{-[(-\ln\rho_1)m_2 + (-\ln\rho_2)n_2]^{1/2}\} \qquad \text{(Eq. 6.2)}$$

Where:

$R(m,n)$	is the 2-dimensional (spatial) autocorrelation function;
ρ	is the correlation coefficient between adjacent pixels ("one-step correlation");
σ^2	is the variance of the random field;
ln	is the natural logarithm ($\log_e$);
m,n	are the horizontal and vertical coordinates of the digital image which correspond to the subscripts 1 and 2 of the one-step correlations.

The separable form of the ACF is more mathematically tractable but is usually less realistic than the nonseparable form. The principal feature of both forms of ACF is that the spatial correlation falls off exponentially with distance between the pixels. Thus, the form of the ACF is simply a symmetrical exponentially declining function where the rate of decline with distance is given by ρ (the degree of statistical dependence between the constituent pixels or "one-step correlation").

Having established a model of the statistical properties of the temporal noise and the nature of its ACF it is now possible to develop the processor to decorrelate the noise prior to the detection stage - the "whitening filter" or decorrelator. However, before the development can be usefully undertaken it is necessary to validate the theoretical model of the temporal noise by comparing it with real data derived from a sample of accurately registered radiographs of growing bone. This is addressed in the following chapter.

CHAPTER
7

EXPERIMENTAL INVESTIGATIONS - MATERIALS AND METHODS

7.1 INTRODUCTION

The purpose of this study was to design and implement a processor to detect stable reference sites in growing bone. The processor has been designed on the basis of theory - the mathematical model and its analysis - but the processor must function in the practical world. Consequently, an experimental approach is needed not only to test the agreement between the model and actual data but also to test the processor to see if it performs well enough to be useful in practice. An additional reason for supplementing theory by experiment is the need to estimate the parameters of the temporal noise model to permit the noise to be adequately decorrelated by the decorrelation filter.

7.2 EXPERIMENTAL DESIGN

To fulfil the objectives of validation, estimation and testing, requires data drawn from serial radiographs of growing bone where the true position of registration is already known. At present, such data can only be provided by controlled projection radiography combined with implanted radiopaque tantalum markers. The acquisition of radiographs fulfilling these requirements poses several ethical and practical problems which have severely restricted the availability of such samples. With the increasing limitations imposed on the use of implanted markers and serial radiography it is now unlikely that large samples will ever become available (Mathews and Payne, 1982; Poswillo, 1984). Consequently, the objectives must be achieved using the small sample which has become available for study.

However, if the entire sample is used for each of the procedures (that is, to validate the model, estimate the parameters and test the resulting processor) the ultimate outcome should be a processor that works well on those particular data, because it has been designed, refined and finally tested on one set of data. This single sample approach is referred to as "testing on the training data" and is known to lead to over-optimistic estimates of performance (Duda and Hart, 1973).

One of the most important aspects of processor design is to ensure that the performance is maintained when the test sample is somewhat different to the sample used for design: that is, the processor must be robust. If the same data are used for designing and testing we have no clear idea, other than from theory, how robust the processor is because an optimistic assessment will result.

To avoid this problem it is necessary to have separate, statistically independent data to perform these tasks. Where only small samples of accurately registered data are ever likely to be available this poses a dilemma because the existing sample must be partitioned in some way to provide independent samples for the different tasks. Partitioning the data into smaller samples reduces the efficiency with which the model parameters may be estimated and also reduces the precision and power of any statistical tests which are applied to the performance of the processor. In addition, the way in which the data is divided poses a further problem. If most of the data are reserved for validation and estimation there can be little confidence in the test of the processor; if most of the data are reserved for testing the processor we cannot have confidence in the design.

There are several possible solutions to this problem. The one which has been adopted in this study broadly follows that described by Fukunaga (1990) for the design of automatic pattern classifiers where only limited data are available.

The procedure consists of three parts:

i. The data are partitioned into two, approximately equal, subsets.
ii. Model validation and parameter estimation are performed using one subset.
iii. Performance of processor is tested separately on <u>both</u> subsets.

1.7.3 EXPERIMENTAL MATERIALS - RADIOGRAPHIC DATA SETS.

Three primary data sets, in the form of radiographic images were employed in this study. These were:

Data set 1.

Paired serial lateral cephalometric radiographs of 50 children with implanted tantalum markers in the upper and lower jaws drawn from the human implant research files of the University of Washington, Seattle, USA (Figure 7.1). The subjects covered the full range of malocclusion from class III to class II, however, no pathological cases were included. The details of the cephalostat and radiographic method may be found in Julius (1972). The original radiographs were duplicated on single emulsion radiographic duplicating film (Kodak X-Omat duplicating film, Kodak Ltd., Hemel Hempstead, Herts.) prior to analysis. The pairs of duplicated films were divided into three subsets.

Figure 7.1 Example of Cephalometric Lateral Skull Radiograph from University of Washington implant sample.

First, the films of two subjects were selected to provide one subject with clear images and well defined bony details; and one subject with poor images and ill defined bony details. These films were used for prototyping of the computer algorithms and development of the processor. The remaining forty-eight pairs of films were allocated to one of two stratified random samples of 24 pairs each (Table 7.1). To obtain the samples the subjects were stratified by: (i) age at initial radiograph; and (ii) period between radiographs to provide, 6 strata. Each stratum was then divided using random permuted blocks within the strata (Pocock, 1983). No distinction was made between orthodontically treated and untreated individuals.

Table 7.1. University of Washington serial cephalometric lateral skull radiographs.

Data set number	Number of film pairs	age at initial film in years		time between films in years		Male/female ratio
		(medians and inter-quartile ranges)				
1.	50	11.69	11.08-12.71	4.04	3.25-4.71	19/31
1a.	24	11.66	10.99-12.49	4.01	3.16-4.69	8/16
1b.	24	11.79	11.02-12.51	4.20	3.09-4.86	11/13

Data set 2.

Comprised a further group of serial lateral cephalometric radiographs of 4 children (2 male, 2 female) with implanted tantalum markers in the upper and lower jaws drawn from the human growth study ("The Mathews' - UCSF Sample") of the University of California, San Francisco, California, USA (Baumrind et al., 1987). The subjects had mild class II malocclusions. Three serial radiographs of each subject, covering a period of eight years around the time of puberty, were selected on the basis of a clear anatomical image. The original radiographs were duplicated on single emulsion photographic film at the Berkeley Laboratory, Department of Growth and Development, UCSF. An optical density step-wedge and fiducial marks were added to the margins of each film during the duplicating process. The magnification factor at the mid-sagittal plane (MSP) was constant at 1.08 for all films in data sets 1 and 2.

Table 7.2 University of California (UCSF) serial cephalometric lateral skull radiographs.

subject number	number of films	age at initial film (years)	age at final film (years)	sex
1	3	8.67	15.08	f
2	3	8.58	15.16	f
3	3	8.83	15.08	m
4	3	8.92	15.00	m

Data set 3.

This comprised lateral cephalometric radiographs of two dried and defatted adult human skulls with bilateral metallic markers secured to the external surfaces of mandible. The skulls were of unknown origin commercially prepared for teaching human osteology. Multiple radiographs were made of both skulls using the cephalometric unit, Eastman Dental Hospital, London.

7.3.2 Stability of the implanted markers in data sets 1 and 2

For each subject in data sets 1 and 2 the relative position of the implants was examined for to verify their stability (to the limits attainable for single two-dimensional films; Rune, et al., 1980) before inclusion in the study. The procedure employed to verify the stability of the implants is described in Appendix B.

7.3.1 Nomenclature for images from data sets 1 and 2.

Throughout the study the images derived from data sets 1 and 2 were examined in serial pairs. The digital images acquired from the earlier and later radiographs of each pair will be referred to as: the *earlier image* and the *later image* respectively. The image representing the difference between correctly registered pairs of radiographs will be referred to as the *difference image.*

7.4 EXPERIMENTAL APPARATUS AND METHODS.

7.4.1 **Experimental Apparatus**

The major experimental investigations, and practical implementation of the detection processor, were carried out using a digital image processing system which consisted of three major components: the imaging system; digitizers and image storage; and digital computers to process the image data.

7.4.1.1 *Hardware.*

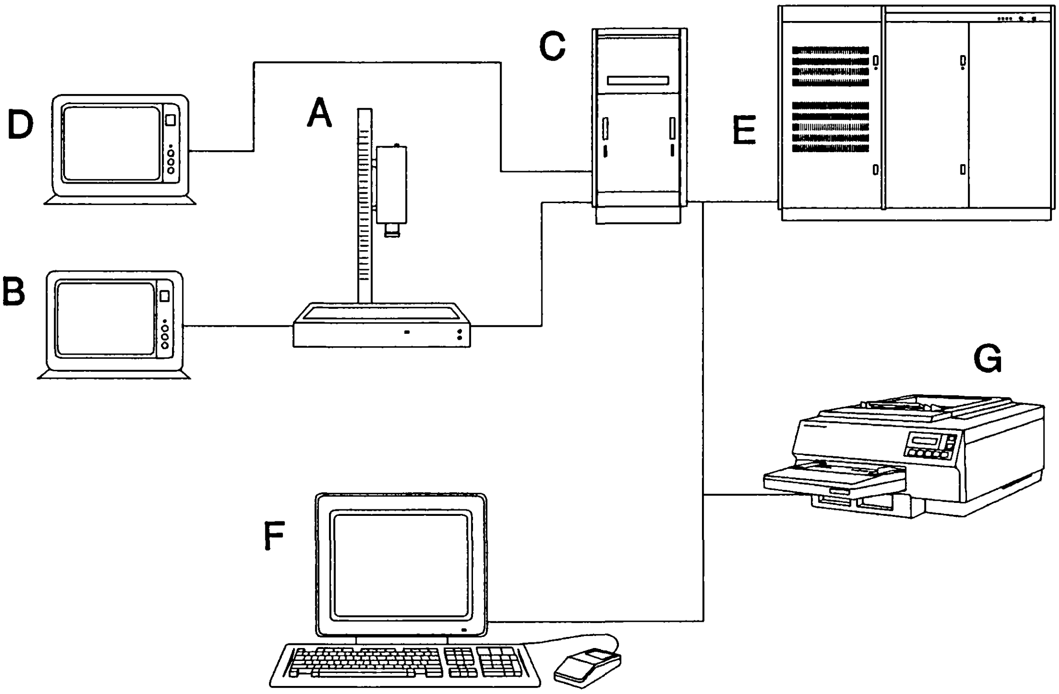
The components of the image processing system employed in the present study are shown diagrammatically in Figure 7.2. The image acquisition was performed by a high quality high-resolution Vidicon camera(2) with a macro lens suspended above a variable intensity fluorescent light box. The camera was linked directly to a video monitor (National Panasonic WV-5310)(3) which allowed focusing and fine adjustment of the position of the radiographs beneath the camera. The camera output was also fed to a CRS-4000 frame grabber/framestore(4) where the electronic signal from the camera was sampled and stored. A second video monitor (National Panasonic WV-5410)(3) displayed the sampled image.

The sampled image was passed to the main computer (Sun 3/160)(5) where the processing operations were performed. Printed copies of raw and processed images were provided by a high resolution laser printer (LaserWriter II)(6).

7.4.2 **Control of the computer-based image processing system.**

The entire system of image processing devices was controlled by a workstation linked to the main computer which employed a UNIX system V operating system(7).

To facilitate the use of different hardware devices and to allow processing operations to be specified independent of the particular device which was performing the task, the overall operation of the system was controlled using the device independent image processing library (DIMPL) computer based routines (Greenwood and Javid, 1991). This library of processing routines and image processing operations is controlled by the high-level computer language "YAM" (Greenwood and Javid, 1989; Javid and Greenwood, 1991).



A Vidicon camera suspended above a variable intensity light box (for trans-illumination of radiographs);
B video monitor coupled directly to the camera;
C CARS-4000 frame store buffer (where the image is sampled and digitized);
D video monitor which displays the sampled and grabbed image;
E main computer (housed in a separate room);
F image processing computer workstation;
G high resolution laserprinter.

Figure 7.2 Diagrammatic representation of the experimental apparatus in the image processing laboratory.

YAM is an interpreted language which employs source codes written in the computer language "C" to perform the basic image processing tasks (Kerningan and Ritchie, 1978). YAM operates as an interactive programming tool providing fast interaction for computer algorithm development and for prototyping solutions to image processing problems. A graphical interface known as YAMTOOL provides access to the YAM environment and provides all the basic image acquisition, processing and output operations in a simple format controlled by a computer "mouse". However, for more specialised processing operations and for applications which are computationally complex, it is possible to switch to software programs written directly in "C" which are outside the YAM environment. This has the advantage of reducing the execution time for the complicated tasks and allows access to processing methods which are not possible with YAM.

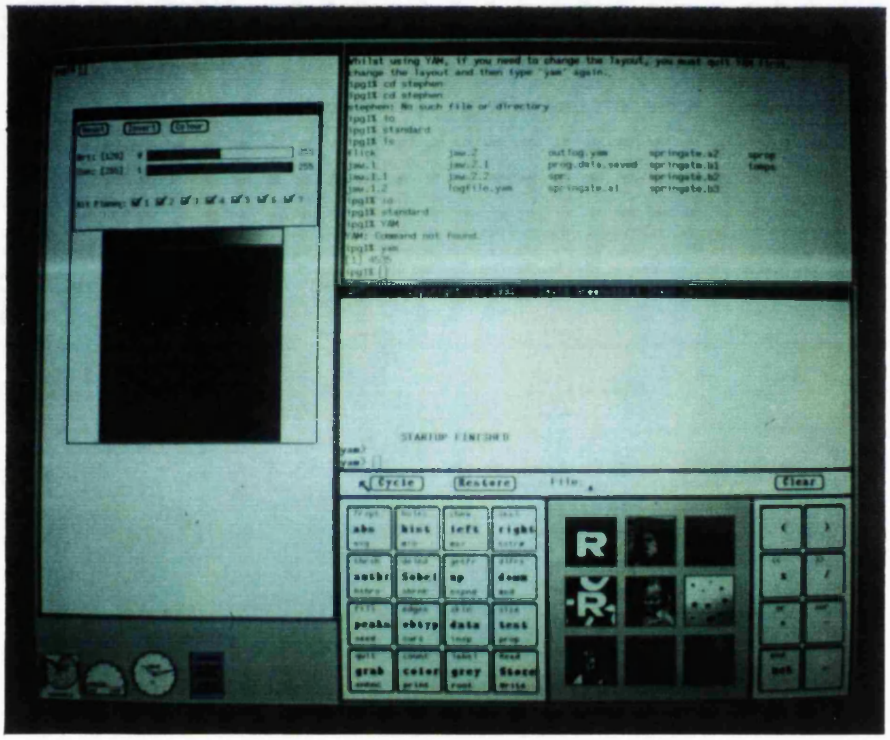
7.4.3 Image processing operations using YAM

To process an image using YAMTOOL the image must be displayed in the YAM processing window. This is simply an area on the computer screen where the active image is displayed. Images can be called into the main YAM processing window (on the computer monitor screen) directly from the video camera and framestore, from temporary storage or alternatively images can be recalled from permanent storage. Once in the processing window images can be processed using one of 45 basic operations.

The operations available in YAM consist of arithmetic, logical (Boolean), as well as standard neighbourhood, comparative and interactive functions. A full description of the available functions is given in Javid and Greenwood (1991). The results of processing can be placed in temporary storage or transferred directly to permanent storage in the tape facility of the main computer. The YAMTOOL interface and Unix processing windows as they appear on the computer screen are shown in Figure 7.3.

7.4.4 Image processing operations outside YAM

In order to process YAM images outside YAM conversion of the digital image to the appropriate format was required. Two other formats were used in this study: *pixrect* and a *raw format* (in which the YAM image header was removed or "stripped" from the data file). The conversion between YAM and pixrect or raw format was carried out using the appropriate commands written in the computer programming language "C".



```
Whilst using YAM, if you need to change the layout, you must quit YAM first,
change the layout and then type 'yam' again.
tpg1% cd stephen
tpg1% cd stephen
stephen: No such file or directory
tpg1% cd
tpg1% standard
tpg1% ls
#lick           jaw.2           outlog.yam       springate.a2    sprep
jaw.1           jaw.2.1         prog.data.saved  springate.b1    temps
jaw.1.1         jaw.2.2         spr              springate.b2
jaw.1.2         logfile.yam     springate.a1     springate.b3
tpg1% cd
tpg1% standard
tpg1% YAM
YAM: Command not found.
tpg1% yam
[1] 4535
tpg1% 
```

```
STARTUP FINISHED
yam>
yam> 
```

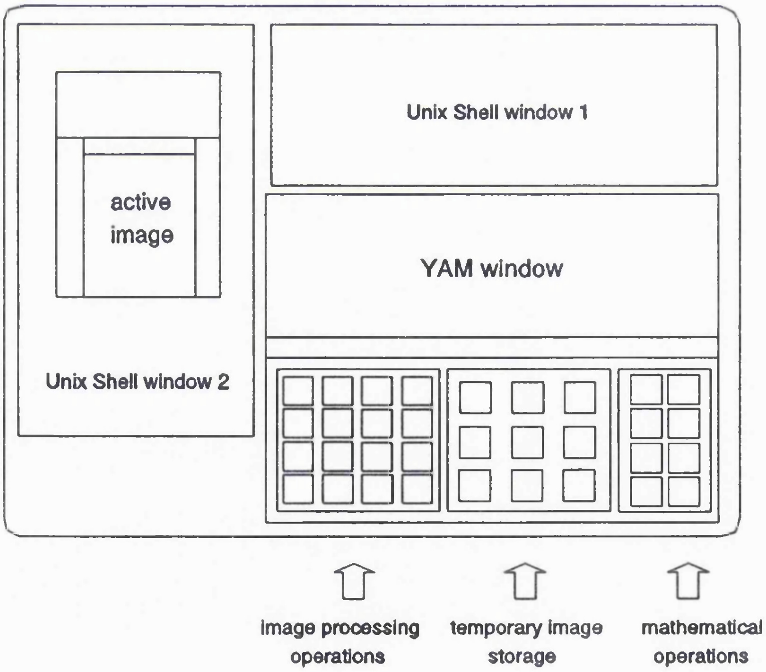


Figure 7.3 YAMTOOL interface and Unix processing windows as they appear on the computer screen.

7.5 PRELIMINARY SET-UP AND INHERENT ERRORS OF THE IMAGE PROCESSING SYSTEM

Several practical image processing procedures were common to all or most of the experimental investigations. The procedures involved calibration and set-up of the apparatus; image acquisition; noise removal; and precise image registration using tantalum markers.

7.5.1 Image resolution

There are three different properties of the digital image embraced in the term resolution: individual pixel size; total image size in pixels; and grey-scale quantization. Each of these was standardized before image processing began.

The video camera and frame grabber provided an eight-bit (256 grey-levels) image which was spatially digitised to provide a maximum resolution of 512 x 512 pixels. However, the speed of many of the processing operations which use YAM can be very slow. To provide the maximum speed for the operations the images used throughout this study have resolutions of either 96 x 96 or 128 x 128 pixels.

7.5.2 Pixel size - sampling interval

The size of the individual pixels is of importance because this determines the size of the smallest structure which can be recorded and displayed. Given the generally poor spatial resolution of radiographic screen-film systems, it was considered important to recover the smallest anatomical details recorded by the films. The maximum spatial resolution for high contrast structures by screen-films systems is approximately 0.2 mm (Barrett and Swindell, 1981a). However, for practical purposes the smallest bony details recorded on screen-film radiographs are approximately 0.28 mm (Ishizuka, 1981). Although the duplication of the radiographs on to single emulsion film would have degraded the resolution of the image, it was decided to recover high contrast details down to ≈0.25 mm. To record detail of this magnitude the pixel size was set at 0.13 x 0.13 mm. This size was dictated by the need to sample the film at twice the maximum spatial frequency to be recovered.

7.5.3 Image capture

Image acquisition or capture was achieved by placing the radiograph on the viewing box beneath a light excluding mask. The Vidicon camera aperture stop was set to f8 and focused. Fine adjustment of the position of the radiograph was made while viewing the computer output on the video monitor. Once the image displayed on the computer screen was satisfactory in size and position, YAMTOOL was instructed to grab and average an appropriate number of images to provide a final image cleaned of electrical line noise. The resulting average image was stored in one of nine temporary storage areas, which remained on display while the workstation was operating.

7.5.4 Spatial quantization - removal of image acquisition noise

All digital images acquired during this study were initially quantized to the full 8 bit depth (256 grey-levels) provided by YAM. The camera and framestore produced electrical line noise which was superimposed on the image signal. Electrical line noise mainly contaminates the least significant bitplanes of the image signal, thereby reducing the number of grey-levels which contain useful information. The level of electrical line noise could be reduced by minimising the gain applied to the camera signal. In all cases the camera was operated with the gain control off. Nonetheless, this noise was still quite considerable. However, because the noise was zero-mean and white, its effect could be reduced by averaging several images or "frames" recorded by the camera over a short period of time. It was decided that the image signal should have 128 clean levels, that is, the noise should be limited to the least significant bitplane, which could be discarded. The number of frames to be summed and averaged was determined by operating the camera with the light-tight lens cover in place. Blocks of 8 frames were grabbed, summed and averaged. Consecutive blocks were subtracted to provide a difference image, which was stored. This process was repeated to produce a series of images which represented the difference between multiples of 8 frames. The individual bitplanes for each difference image were then displayed and the pixel values examined. After 32 frames were averaged it was found that structure existed only in the two lowest bitplanes. After 48 frames were averaged the noise was restricted to 5% of the pixels in the second lowest bitplane and was not detectable in higher planes. Since all but a small proportion of the noise was confined to the lowest bitplane and because of the time consuming nature of frame averaging it was decided to limit the averaging to 48 frames and accept the small residual noise.

To reduce this source of random variation between images during processing and analysis this bitplane was replaced with a plane of zeros. This was accomplished by shifting the image planes down through the bit stack by one plane followed by inclusion of a new zero-value plane at the lowest level in the bit stack. Consequently the images used in this study were acquired with full 8 bit quantization depth but were processed with only 7 bits; thereby providing 128 grey levels.

7.5.5 Low frequency noise

The intensity of illumination across the image was examined for inconsistencies by measuring the mean intensity of the image obtained from the light box and from the camera alone. In three separate tests a spatially varying mean intensity was found across the image in both planes of space (Figure 7.4) which persisted even when a light tight cap was fitted to the camera lens. This problem was rectified by avoiding continuous operation of the camera and by examining the mean illumination at the beginning and end of each experimental session.

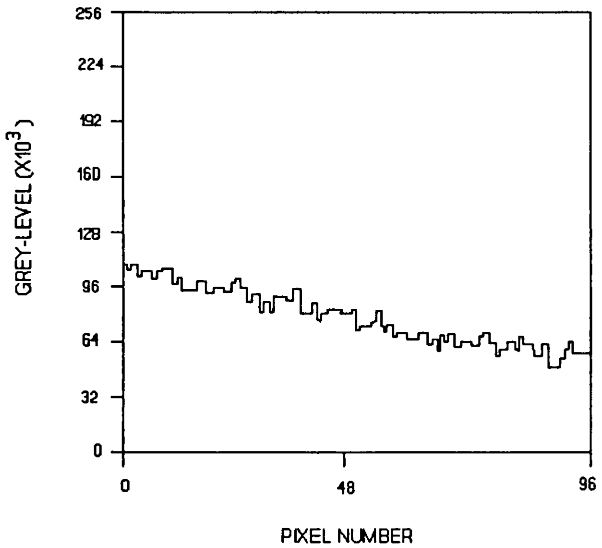


Figure 7.4 Graph showing spatially varying mean intensity across the camera face after prolonged operation.

7.5.6 Geometric fidelity of the images

Vidicon cameras are subject to several sources of distortion affecting the geometric fidelity of the image.

The geometric fidelity of the camera image was examined using a precision graticule(8), ruled into divisions of 0.1 mm over 1 cm^2, placed at the approximate centre of the camera field. The image was acquired at a pixel size of 0.1 mm. The resultant image was examined for inconsistencies in the number of pixels between the divisions. The central region of the camera field was found to record an image with variations of less than one pixel per 90 pixels.

7.5.7 Registration of serial images

The majority of the experimental investigations required the serial images to be registered precisely by accurate alignment of the implant images. This was achieved using pairs of serial images in a two stage procedure.

7.5.7.1 *Coarse alignment of the images*

The positions of corresponding implants were located in each pair of radiographs by eye. Having set up the viewing-box and camera the later image was placed on the viewing box and adjusted to bring the area of interest to the central region of the frame grabber. The lateral margins of the radiographs and viewing-box were masked using a black card. The mask and radiograph were secured in place with a series of lead weights. The positions of the implants were then recorded on the camera video monitor using movable/adhesive markers. Multiple images were grabbed, averaged and stored in temporary memory. The radiograph was removed and replaced with the earlier radiograph using the markers on the monitor screen to achieve alignment. Images of the earlier radiograph were then grabbed, averaged and stored in temporary memory.

7.5.7.2 *Fine registration of the images*

The coarse registration of the images generally produced an alignment within two or three pixels of the correct position (point at which the implants achieve maximum match). The final fine registration employed computer-based matching of the implant images using a minimum-variance histogram method (Green et al., 1978).

The earlier and later images were transformed to approximately the same mean and variance using YAMTOOL functions. The transformed images were then aligned by locating the position of minimum difference between the two images of each implant. This was done by removing all detail from the images except a square approximately 2 x 2 mm around the image of one implant. The earlier image was shifted by individual pixels in both x and y planes until the variance of the difference-histogram reached a minimum. This process was repeated for each implant in turn. The shifts necessary to achieve fine alignment were then used to adjust the position of the earlier image by pixel increments. This image and the later image were then transferred to permanent storage.

7.5.7.3 *Registration of regions without implants.*

Where registration was required for regions situated away from the implants it was not possible to achieve fine alignment of the images at the chosen resolution, as described above. In this case the images were registered using a larger format (e.g. 256 x 256 pixels) but with the same pixel size (0.13 mm); thus each image covered a larger anatomical area. Once satisfactory alignment had been achieved the image was reduced to 96 x 96 format by removing the appropriate number of columns and rows of pixels.

7.5.8 **The anatomical sites investigated - subimage blocks.**

The images were captured from the radiographs in 1.55 cm^2 blocks from five anatomical sites in the mandible and two sites in the maxillae (Figure 7.5) avoiding, as far as possible, the roots of the teeth. This allowed the examination of consistent anatomical regions and facilitated data handling while ensuring accurate control over the spatial resolution of the images.

Four regions of the jaws were not included in the anatomical coverage: in the mandible, the most superior aspect of the ascending ramus and the most infero-posterior part of the mandibular angle were not imaged; and in the maxilla, the region of the anterior nasal spine and maxillary tuberosity were not included. These regions were not included in the study because remodelling of the bone eliminated the original sites in some or all of the subjects.

To avoid damaging the delicate single emulsion layer on the duplicate radiographs the films were orientated facing to the left. This orientation was adopted for all lateral cephalometric images used in the study.

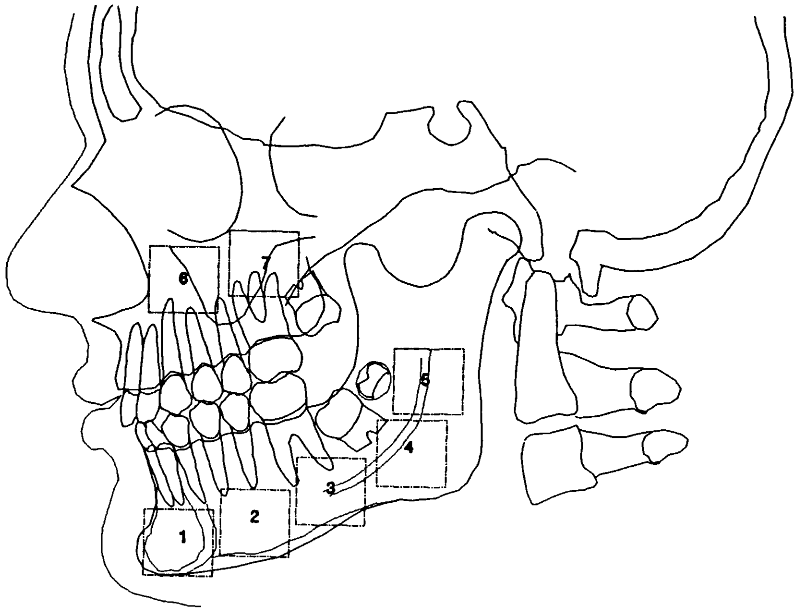


Figure 7.5 Approximate positions of the image blocks examined in the experimental investigations.

7.6 DATA HANDLING AND STATISTICAL ANALYSIS

7.6.1 Handling of digital image data

The raw data obtained in this study were generally in the form of digital images. The analysis and handling of data in digital image form poses problems which are not apparent in conventional biomedical analysis. Specifically, each pixel in an image is considered as a data point in a two-dimensional matrix. Thus, each 96 x 96 pixel subimage required the recording and processing of almost 10,000 data points and in total it was necessary to examine over 4,000,000 data points. Consequently, the standard methods of manually tabulating data into a spread-sheet prior to analysis would not have been possible. Instead, basic descriptive and summary statistics were derived using the *HIPS*-based software programs(9) on the Sun 3/160 computer, using data read directly from the image files. These basic descriptive and summary statistics were then saved as a "hard copy" for subsequent use in statistical analysis programs not requiring access to the raw data.

Where statistical analysis of specific raw (pixel) data was required, the pixel parameters (spatial coordinates and grey-level value) were read into the *outlog.yam file* using the YAMTOOL cursor and inspect functions, which were then saved and tabulated into ASCII data files or SPSS/PC+ files for comparative statistical analysis using standard statistical programs.

7.6.2 Statistical analysis

The experimental data accumulated in the various investigations were analysed using four computer programs:

SPSS system for statistical analysis(10) (SPSS Inc. 444 N. Michigan Avenue, Chicago, USA);

Simstat statistical analysis program(11) version 2.1 (Provalis Research, Montreal, Canada);

Arcus professional analysis package(12) version 1.5 (Medical Computing, Aughton, Lancashire, UK);

Plot3d(13) (Sun Microsystems Europe Inc., Camberley, Surrey, UK).

Plot3d was implemented on a Sun 3/160 computer(5); SPSS was implemented on an IBM (model 70) 386 PS/2 computer(14); and the remaining programs were implemented on an IBM compatible Vig III/33 personal computer(15).

7.6.3 Statistical methods

7.6.3.2 *Preliminary analysis*

Preliminary analysis was aimed at ensuring the data conformed to the assumptions underlying the proposed statistical tests. The normality of data sequences was tested using *Shapiro-Wilk's W* test (Royston, 1982). Where the Shapiro-Wilk W statistic attained significance at $p \leq 0.05$ the sample was deemed not to be a random sample from a normally distributed population.

7.6.3.3 *Descriptive Statistics*

For data which followed a normal (Gaussian) distribution standard descriptive statistics of sample mean, standard deviation, variance, standard error of the mean, and 95% confidence interval were used. Where data were not normally distributed the median, inter-quartile range and 90-percentile confidence interval were used. The descriptive second-order statistics (autocorrelation) for two-dimensional samples are given as correlation coefficients (ρ) regardless of the underlying distribution.

7.6.3.4 *Comparative Statistics*

Where data were normally distributed the variances of two samples were compared using *Snedecor's variance ratio* test; the variances of multiple equal sized samples were compared using *Hartley's* F_{max} test and for unequal sized samples *Cochran's* test for the homogeneity of variances was used. Where the variances did not differ significantly, comparative analysis of two samples was conducted using two-tailed paired or unpaired *t* tests.

Where data were not consistent with a normal distribution, comparisons between paired samples was conducted using *Wilcoxon matched pairs signed ranks* test. Comparisons between multiple unrelated samples were conducted using the *Kruskal-Wallis* test together with a multiple comparison of means to locate the sites of statistical significance. Differences in frequencies (proportions) between groups were tested using χ^2 analysis or, where the expected values were less than 5, *Fisher's exact probability* test was used instead.

Dependence between pairs of variables was assessed visually to exclude non-linear association followed, where appropriate, by quantitative analysis. Where data were consistent with an underlying bivariate normal distribution the association between them was assessed using *Pearson's product-moment correlation.* Where the data were not consistent with a bivariate normal distribution, the dependence between variables was assessed using *Spearman's rank correlation* test.

CHAPTER
8

MODEL VALIDATION: COMPARISON OF THE MODELS WITH DATA.

8.1 INTRODUCTION

In the previous chapters two theoretical models have been formulated. The first, is a general model based on the classical model for the detection of a signal submerged in noise. The second model is a more detailed description of the features which hinder detection - the temporal noise.

The processor has been derived on the basis of these models and in order to have confidence in the practical application of the processor it is necessary to confirm the agreement between the models and real data on which it is to operate. This process is usually referred to as *model validation.*

8.1.1 The model assumptions to be validated

In order to validate the two theoretical models the assumptions which constitute their mathematical structures must be examined and tested.

The general model embodied in equation 5.1 is represented by three assumptions:

(i) signal is deterministic and known;
(ii) signal and temporal noise are combined additively;

(iii) statistical independence of signal and temporal noise.

An additional assumption is required for the detection of one signal from M possible signals (*M'ary* detection):

(iv) signals (subimages) are statistically stationary.

The temporal noise model incorporates a further three assumptions about the statistical structure of the temporal noise:

(v) Gaussian PDF with a mean of zero
(vi) statistically stationary
(vii) Exponentially declining ACF (first-order Markov)

These seven assumptions represent the mathematical features which are important for analysis of the models. However, assumptions (i) and (ii) have already been imposed *a priori*, by defining the image in the initial radiograph as the signal; and the temporal noise as the arithmetic difference between correctly registered earlier and later radiographs. Accordingly, it is only necessary to test the validity of assumptions (iii) to (vii).

8.2 MATERIALS AND METHODS

The assumptions were tested by comparison of their theoretical forms with data derived from pairs of serial lateral cephalometric radiographs of the mandible and maxillae derived from the twenty-four subjects in Data set 1a.

For each subject, the differences between the theory and data were subjected to statistical hypothesis testing except for assumption (vi) which was not submitted to any formal statistical test. The critical value for each test statistic was set at $p=0.05$ to provide a exacting test of the null hypothesis. Where an assumption was subjected to hypothesis testing, the proportion of subjects for whom the null hypothesis was rejected was tested against the expected proportion using *Fisher's exact probability* test.

Throughout the validation the assumptions were tested only within each individual and no attempt was made to combine the data from the twenty-four subjects prior to analysis. The mandibular and maxillary images were analyzed separately for each subject.

8.2.1 Image capture and preliminary processing

Images from the earlier and later radiographs of each pair were captured as a series of subimages (approximately 1.55 cm^2) of 96 x 96 pixels with a pixel size of 0.13 mm from five anatomical sites in the mandible and two sites in the maxillae as described in section 7.5.8 (see Figure 7.5). Corresponding subimages from each pair of radiographs were adjusted to a similar mean and variance, and registered using the tantalum markers as described in section 7.5.7. The seven subimages from the earlier radiograph were then subtracted from the corresponding regions in the later image and the difference image stored.

The testing of assumptions (iii), (iv) and (vi) required the estimation of one or more image properties which are defined over a neighbourhood rather than for individual pixels. However, the size of the neighbourhood was known to influence the properties on which the validations were to be based (Hunt, 1981). It was important, therefore, to use a similar neighbourhood size to that proposed for the final processor. Consequently, a square neighbourhood 15 x 15 pixels (1.95 mm x 1.95 mm) was selected.

Four 15 x 15 pixel samples were drawn from the mandibular subimages and 10 from the maxillary subimages of each earlier image and from corresponding locations in each associated difference image. The spatial positions of the samples were determined randomly within each subimage.

8.2.2 Testing the model assumptions

8.2.2.1 *Assumption (iii): statistical independence of signal and temporal noise*

The practical significance of statistical dependence between signal and noise lies in the possible correlation between the variance or amplitude of the noise and the mean value of the signal. This possible association was investigated by examining the relationship between the means of the earlier images and the variances of the difference images.

The means of the 15 x 15 pixel samples from each earlier image were plotted against the variance of the corresponding sample from the difference image. The plots were first examined visually to assess the extent of any non-linear association between the variables followed by a quantitative examination of the strength of the linear association using *Spearman's correlation coefficient*. The statistical significance of the correlation coefficient was used to test the hypothesis: the correlation between the mean of the earlier image and the variance of the difference image is zero.

8.2.2.2 *Assumptions (iv) and (vi): statistically stationary signal and temporal noise*

The statistical stationarity for the signal was examined by testing the homogeneity of the variances between the twenty 15 x 15 pixel samples of the earlier image within each jaw. The hypothesis that the variances were equal was tested using *Hartley's F-max* test.

The statistical stationarity of the temporal noise was tested in exactly the same way but using the samples drawn from the difference image rather than the earlier image.

8.2.2.3 *Assumption (v): the temporal noise has a Gaussian PDF*

The temporal noise is assumed to be a sample from a Gaussian distribution with a mean of zero. This assumption was tested by examining the distribution of the grey-levels in the difference image.

The distribution of pixel values was examined qualitatively for each (96 x 96 pixel) subimage of the difference image using the YAMTOOL histogram function which calculated and displayed the grey-level histogram. Each histogram was examined visually to see if it conformed to an approximate Gaussian distribution. In addition, a more detailed quantitative analysis was performed using a sample of 100 pixels drawn from randomly determined sites across the subimages for each jaw. The hypothesis that the sample was a drawn from a Gaussian population was tested using the *Shapiro-Wilk's* test. However, the hypothesis that the mean of the distribution was zero could not be validly tested because the images had been adjusted to give similar mean values. Thus, the difference image would be expected to have a mean of approximately zero irrespective of the validity of this assumption.

8.2.2.4 *Assumption (vii): exponentially declining ACF*

The validation of this assumption required the estimation of the ACF from large image regions greater than 30^2 pixels to provide sufficient resolution. To estimate the ACF using such a large data set for all subimages of each test subject was impracticable, requiring about 100 hours of continuous operation of the Sun 3/160 computer(5). As an alternative, the exponential decline of the ACF with increasing distance was examined using an estimate of the ACF obtained from a single subimage from each subject. The subimages were chosen arbitrarily to provide three estimates from each anatomical region occupied by the seven subimages for 21 of the 24 subjects.

The ACF estimates were calculated for a 31 x 31 pixel sample located arbitrarily in the central 76 x 76 pixels of the subimage. The two-dimensional ACF was calculated for spatial shifts of 0 to 10 pixels at integral displacements. The resulting ACF estimate was plotted as a three-dimensional surface and an arbitrary two-dimensional section was made through the peak (point of zero shift) parallel to the amplitude axis (Figure 8.1).

The correlation values along the surface of the two-dimensional section were plotted on a logarithmic scale against displacement for 0 to 10 pixel shifts. The closeness of the ACF to an exponentially declining form was determined by visual examination of the plotted curve (an exponentially curve appears as a straight line when potted on a logarithmic scale).

8.3 RESULTS

The results are presented in Tables 8.1 - 8.4 and graphically in Figures 8.2 to 8.4. The results of the individual hypothesis tests should be interpreted together with the actual value of the test statistic. The results of the combined hypothesis tests for the whole sample should be viewed as an additional guide to the validity of the assumptions.

8.3.1 Assumption (iii) statistical independence of signal and temporal noise

Preliminary visual analysis of the earlier and difference images revealed associations between them where they depicted endosteal contours, particularly at the lingual border of the mandibular symphysis (Figure 8.2).

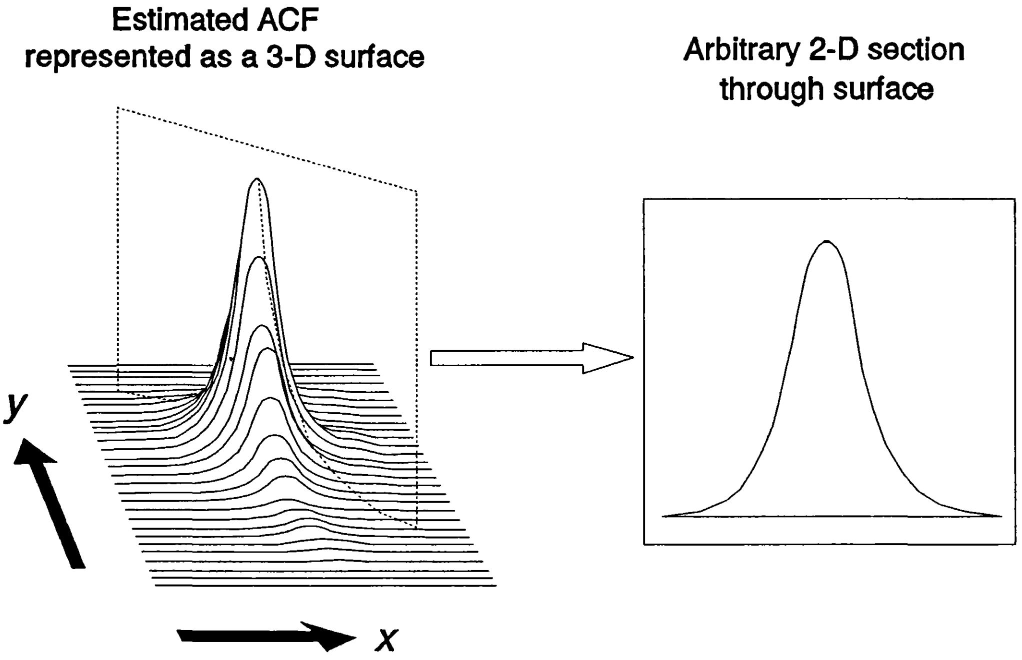


Figure 8.1 Examination of the ACF using a 2-D section through the estimated ACF surface perpendicular to the x,y plane.

In addition, the examination of the graphical scatterplots revealed a slight inverse association between mean of the earlier image and the variance of the difference image. However, although this is supported by a negative Spearman correlation coefficient for 17 of the 24 cases for the mandible and 16 for maxilla, the null hypothesis could not be rejected in 22 of the 24 cases for the mandible and 23 cases for the maxilla (Table 8.1).

The proportion of subjects showing a statistically significant result at $p \leq 0.05$ is close to that expected by chance alone (Fisher's exact test: $p=0.42$ and $p=0.51$, for mandible and maxilla respectively).

8.3.2 Assumptions (iv) and (vi): statistically stationary signal and temporal noise

The results of the tests for the homogeneity of the variance for the earlier image and for difference image are shown in Tables 8.2 and 8.3. For the earlier images the hypothesis of constant variance was rejected in all but three cases for the mandible and

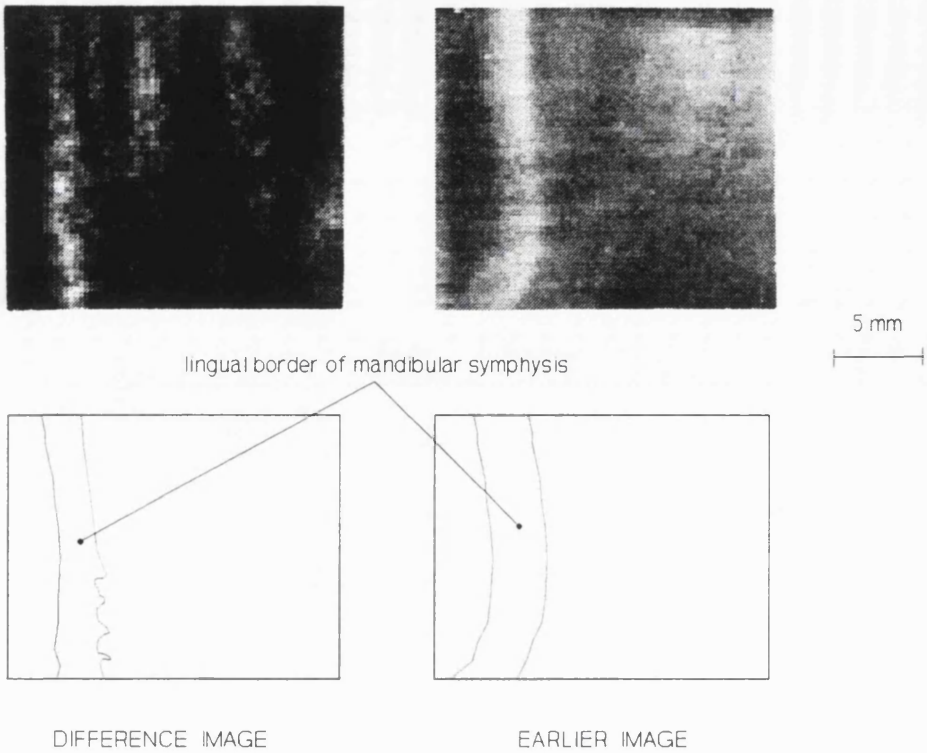


Figure 8.2 – Examples of earlier and difference images showing associations between them at the lingual border of the mandibular symphysis.

four cases for the maxilla. The findings for the difference images were similar, with rejection of the hypothesis of constant variance in all but six cases for both the mandible and maxilla.

Although the critical value of the *F*-max statistic for the rejection of the hypothesis was set at $p \leq 0.05$ the proportion of subjects showing a statistically significant result at $p \leq 0.05$ were very different from that expected by chance alone, both for the earlier image and for the difference image (Fisher's exact test: $p=0.002$ and $p=0.003$, for mandible and maxilla respectively).

Table 8.1 Tests of linear association between mean of the earlier image and variance of the difference image.

subject number	Mandible Spearman correlation coefficient (R_s)	Maxilla Spearman correlation coefficient (R_s)
1	-0.15 ns	-0.18 ns
2	-0.23 ns	-0.20 ns
3	0.02 ns	0.06 ns
4	-0.35 ns	-0.41 *
5	-0.16 ns	-0.29 ns
6	-0.08 ns	-0.31 ns
7	-0.41 *	-0.19 ns
8	0.23 ns	0.16 ns
9	-0.04 ns	-0.24 ns
10	-0.14 ns	-0.19 ns
11	-0.25 ns	0.18 ns
12	-0.24 ns	-0.21 ns
13	-0.11 ns	-0.26 ns
14	-0.13 ns	0.12 ns
15	0.15 ns	-0.10 ns
16	0.11 ns	-0.24 ns
17	-0.39 *	0.09 ns
18	0.24 ns	-0.14 ns
19	-0.12 ns	-0.22 ns
20	-0.15 ns	0.11 ns
21	0.17 ns	-0.21 ns
22	-0.24 ns	0.19 ns
23	0.19 ns	-0.27 ns
24	-0.21 ns	0.05 ns

Notes: 1. Twenty pairs of samples for each subject.
2. The variables examined were derived from 15 x 15 pixel samples drawn from corresponding locations in the earlier and difference noise) images.
3. The levels of statistical significance are indicated thus:
ns $p > 0.05$;
* $p \leq 0.05$

Table 8.2 Test of statistical stationarity for the earlier image.

	Mandible		Maxilla	
subject number	*F*-max statistic	stationarity rejected ($p \leq 0.05$)	*F*-max statistic	stationarity rejected ($p \leq 0.05$)
1	2.23	Yes	2.14	Yes
2	2.56	Yes	2.29	Yes
3	1.68	Yes	2.51	Yes
4	1.84	Yes	1.69	Yes
5	1.76	Yes	3.48	Yes
6	1.71	Yes	1.13	No
7	2.60	Yes	2.07	Yes
8	1.94	Yes	1.99	Yes
9	2.48	Yes	2.37	Yes
10	1.73	Yes	1.10	No
11	1.35	No	1.83	Yes
12	2.02	Yes	1.77	Yes
13	1.77	Yes	1.38	No
14	1.82	Yes	1.68	Yes
15	1.80	Yes	2.21	Yes
16	1.69	Yes	3.06	Yes
17	1.21	No	1.73	Yes
18	2.14	Yes	2.48	Yes
19	3.30	Yes	1.61	No
20	2.42	Yes	1.76	Yes
21	2.25	Yes	3.28	Yes
22	1.31	No	2.64	Yes
23	2.15	Yes	2.12	Yes
24	1.83	Yes	1.99	Yes

Notes.

1. Sample size 15 x 15 pixels, drawn from 20 locations in each jaw.
2. *F*-max calculated for 19 and 224 degrees of freedom.

Table 8.3 Test of statistical stationarity for the difference image.

	Mandible		Maxilla	
subject number	*F*-max statistic	stationarity rejected ($p \leq 0.05$)	*F*-max statistic	stationarity rejected ($p \leq 0.05$)
1	1.75	Yes	2.44	Yes
2	2.39	Yes	1.03	No
3	2.48	Yes	2.01	Yes
4	2.32	Yes	1.74	Yes
5	1.56	No	2.48	Yes
6	1.94	Yes	2.29	Yes
7	1.31	No	2.07	Yes
8	1.71	Yes	1.99	Yes
9	1.95	Yes	1.57	No
10	1.09	No	2.10	Yes
11	1.85	Yes	2.33	Yes
12	1.02	No	1.69	Yes
13	1.77	Yes	1.98	Yes
14	1.82	Yes	1.62	No
15	1.80	Yes	1.20	No
16	1.99	Yes	1.06	No
17	2.41	Yes	2.63	Yes
18	1.14	No	2.48	Yes
19	2.30	Yes	1.68	Yes
20	1.82	Yes	2.36	Yes
21	2.25	Yes	2.28	Yes
22	1.96	Yes	1.69	Yes
23	1.15	No	1.12	No
24	1.53	Yes	1.39	Yes

Notes.

1. Sample size 15 x 15 pixels, drawn from 20 locations in each jaw.
2. *F*-max calculated for 19 and 224 degrees of freedom.

8.3.3 **Assumption (v): zero-mean Gaussian PDF of the temporal noise**

Preliminary visual analysis of the grey-level histograms revealed a clearly defined unimodal distribution with approximate symmetry about the peak (Figure 8.3). This was supported by a Shapiro-Wilk's statistic which failed to reach significance at the $p \leq 0.05$ in 22 cases for the mandibular image and in 21 cases for the maxillary image (Table 8.4). The proportion of subjects showing a statistically significant result at $p \leq 0.05$ is close to that expected by chance alone (Fisher's exact test: $p=0.42$ and $p=0.30$, for mandible and maxilla respectively).

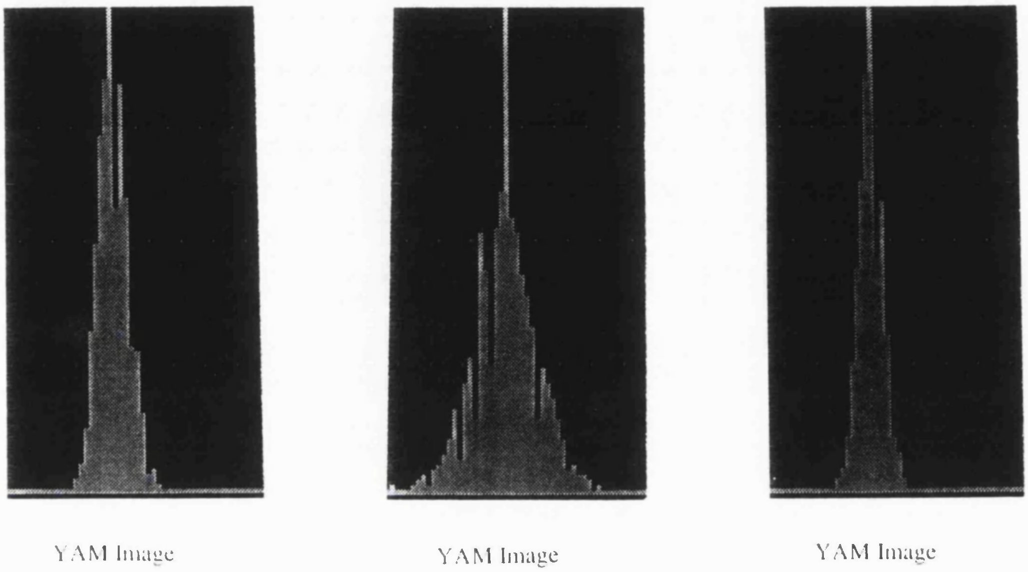


Figure 8.3 Examples of grey-level histograms of difference images showing clearly defined unimodal distributions with approximate symmetry about the peak.

The means of the distributions were calculated in each case and found to be close to zero with a mean for the group of -1.2 (95% CI: -16.5 to 14.1) for the mandible and 2.5 (95% CI: -9.6 to 12.1) for the maxilla.

Table 8.4 Test of Gaussian distribution of the grey-levels in the difference image.

	Mandible		Maxilla	
subject number	Shapiro-Wilk's statistic	Gaussianity rejected ($p<0.05$)	Shapiro-Wilk's statistic	Gaussianity rejected ($p<0.05$)
1	0.988	No	0.980	No
2	0.974	No	0.976	No
3	0.973	No	0.975	No
4	0.984	No	0.981	No
5	0.981	No	0.984	No
6	0.968	No	0.977	No
7	0.971	No	0.982	No
8	0.973	No	0.701	Yes
9	0.984	No	0.980	No
10	0.985	No	0.974	No
11	0.990	No	0.979	No
12	0.979	No	0.846	Yes
13	0.985	No	0.973	No
14	0.982	No	0.978	No
15	0.671	Yes	0.982	No
16	0.981	No	0.985	No
17	0.975	No	0.984	No
18	0.559	Yes	0.976	No
19	0.977	No	0.981	No
20	0.987	No	0.683	Yes
21	0.990	No	0.986	No
22	0.979	No	0.975	No
23	0.978	No	0.979	No
24	0.981	No	0.990	No

Notes.

1. Sample size 100 for each jaw.
2. Hypothesis rejected where Shapiro-Wilk's statistic $p \leq 0.05$.

8.3.4 Assumption (vii): exponentially declining ACF

The closeness of the ACF's to an exponentially declining form was determined by visual examination of the sectioned ACF surfaces alone. No formal quantitative tests were undertaken. Examples of the ACF's are shown in Figure 8.4. The value of the ACF for all subimages reduced with displacement away from the origin. The rate of decline appeared to be consistent with the theoretical exponential form up to displacements of three or four pixels. For larger displacements the estimated ACF declined less rapidly indicating a degree of correlation slightly greater than that predicted by the theoretical model. This finding was true for all the ACF's examined. Thus, for small displacements the ACF's possessed an approximately exponential form although this could not be determined with certainty.

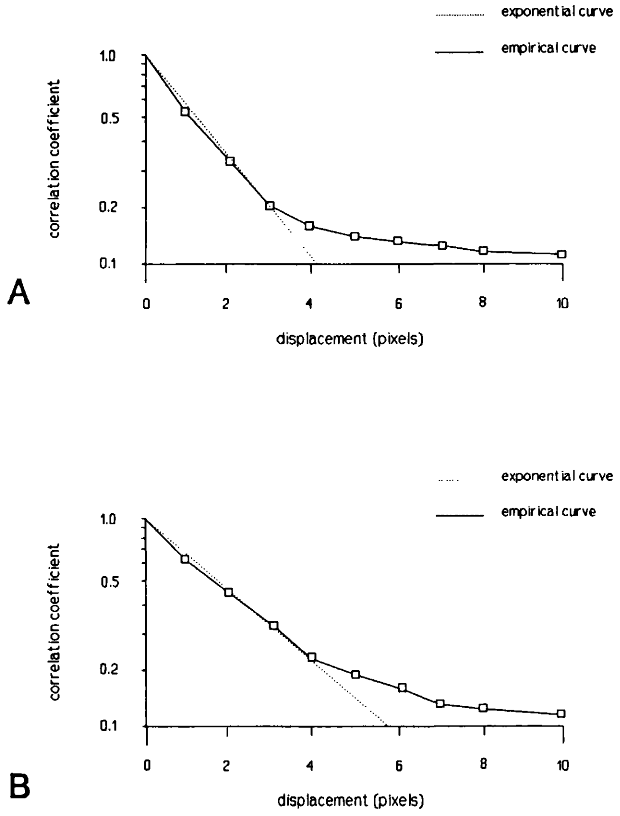


Figure 8.4 Examples of correlation functions for two difference images. The graphs A and B represent half the ACF plotted on a logarithmic scale against displacement in pixels.

8.4 DISCUSSION

Together the two theoretical models form a simple description of the changes in image structure observed in serial radiographs of growing bone. Like most models they portray a process which is simpler than the actual situation they attempt to describe and therefore, they cannot be expected to conform exactly to real data. Consequently, it is important to know what level of agreement is required between theory and actual data for the models to be acceptable. The answer depends on the individual assumptions which constitute the structure of the models. Minor violations of some assumptions may lead to failure of the practical processor, alternatively great differences between theory and reality may have no appreciable influence on the action of the processor.

The findings of this validation study can only be usefully interpreted in relation to the degree of violation which is tolerable for each assumption. This is best seen by examining the reasons for the inclusion of each assumption in the models.

Independence of signal and noise: the significance of statistical dependence between signal and noise lies in the possible correlation between the variance or amplitude of the noise and the mean value of the signal (Cattermole, 1986). When dependence occurs, the noise conveys information in the same way as the signal. To perform detection under these circumstances requires the recovery of this information by an integration operation rather than a correlation operation (Cattermole, 1986). The processor would therefore need to be different to the matched filter derived under the assumption that the signal and noise were independent. However, the practical significance of dependence hinges on the ratio of signal energy to noise energy - the so called signal-to-noise ratio (SNR).

The results of the tests for correlation between the mean values of the signal and the variance of the noise indicate that there is no strong correlation between them. However, the preponderance of negative correlations suggests that, although insufficient evidence was found to reject the null hypothesis, this may indicate a real inverse relationship. This type of association is well recognised in photographic systems (Dainty and Shaw, 1974) and arises from a Poisson distribution of the photons which lead to image formation. The level of dependence between signal and noise is clearly of a low order (if it exists at all) and is unlikely to have any significant effect on the processor except where SNR is low - as envisaged for the images which will be used with the detection processor. Consequently, to mitigate the possible effects of dependence between signal and noise the correlation method used in the final processor will need to be selected from the group of computation methods known to be relatively insensitive to this type of dependence (Secilla et al., 1988).

Statistical stationarity of the signal and temporal noise. The stationarity of the subimages over which the search is conducted is an essential requirement for detection of the correct subimage. The reason is simply that a matched filter detection processor distinguishes between alternative subimages on the basis of image energy which is determined by the mean and variance of the subimages (Ryan and Hunt, 1981). If image energy varies between the subimages, resulting from lack of stationarity of signal or of temporal noise, the detection processor will select the subimage with the greatest energy which may not be the correct match. Consequently, the validity of this assumption is essential to the correct operation of the detection processor.

The tests of these assumptions show them to be violated in almost all cases for the earlier image and in many cases for the difference image. An indication of the magnitude of the violation can be seen by examining the *F-max* statistics which demonstrate a two-fold difference between the variances of the subimages in about half the cases examined. Clearly this magnitude of difference is unacceptable.

The assumption that the temporal noise has a Gaussian PDF with a mean of zero is crucial to several aspects of the design of optimal linear detection filter design. Most importantly it is essential to the interpretation of the matched filter design as an optimal detector (in the sense of minimising the probability of error and minimising the registration error variance; McGillem and Svedlow, 1976). In addition, for the pixel values in an MRF to be derived by a linear combination of the values of its neighbours it is necessary to have a Gaussian PDF (Woods, 1972). However, the practical significance of this is uncertain and, in addition, it is possible to develop a processor without the assumption of Gaussianity (Davenport and Root, 1958). However, such a processor is not a maximum likelihood processor and there is no guarantee that the site selected as the match is in fact the most probable site. Thus, a Gaussian distribution of the temporal noise while not essential is highly desirable.

The results of the tests of this Gaussianity failed to detect a statistically significant difference at $p \leq 0.05$ for the majority of cases, both for the mandibular and maxillary images. Which provides a reasonable degree of confidence that the model is consistent with data examined. However, the means of the distributions were not tested for conformity to the theoretical value of zero. This was because the mean of the difference image was determined by the closeness of the initial histogram matching of the earlier and later images, rather than by the correctness of the theory underlying the temporal noise model. Despite the generally small mean level found for all subjects, the validity of this assumption must remain uncertain.

Exponentially declining ACF: knowledge of the form of the temporal noise ACF is essential for the design of a decorrelator or whitening filter for the temporal noise. The assumption of an exponential form for the ACF allows the decorrelator to have a particularly simple form. However, even quite poor approximations to an exponentially declining form have few practical consequences for image decorrelation (Faugeras and Pratt, 1980). Consequently, even though this assumption was not formally tested it appears that theory provides an adequately close approximation to data.

The findings of this investigation generally indicate good agreement between theory and data for three of the five assumptions with the exception of the assumptions of statistical stationarity of earlier and difference images, which showed poor agreement with data, albeit on the basis of quite stringent tests. The assumption of zero-mean value of the temporal noise could not be validly tested.

The critical importance of statistical stationarity to the operation of the processor indicates that the models cannot be considered valid. There are two options: to reject the models; or attempt to impose the required structure by pre-processing or by transforming the data in some way and thereby compel the data to comply with the assumptions of stationarity and zero-mean.

While for some assumptions it is not possible to impose the required structure, fortunately, statistical stationarity can be imposed by transforming the data from the earlier and later images by the process known as *normalisation* (Ryan and Hunt, 1981). In addition, normalisation provides a zero mean value for the difference image. Unfortunately, these transformations are computationally complicated and greatly increase the time to search for a match (Svedlow et al., 1977). Nevertheless, the importance of statistical stationarity requires that these transformations are adopted if the processor is to work effectively.

CHAPTER
9

IMPLEMENTATION OF THE DETECTION SYSTEM

9.1 STRUCTURE OF THE DETECTION SYSTEM

Before the detection system can be implemented as a practical procedure, the underlying physical operations performed by the system must be determined and organised into a structure which will perform the desired task. This procedure is referred to as *realisation.*

A matched filter detection system can be realised in several ways (Van Trees, 1968). For a two-dimensional system the appropriate structure is known as the *whitening realisation* and this method is the one used as the basis for the practical implementation of the system.

In the whitening realisation the operations performed on the images are decorrelation ("whitening") and cross-correlation. These two operations are specified by the denominator and numerator respectively of the frequency response of the matched filter equation (Eq. 5.1). Both conceptually and practically there are advantages in separating the numerator and denominator of the frequency response to produce two linear processors: the decorrelator; and the cross-correlator. The two processors are then operated in series to produce the same effect as a single matched filter processor.

The decorrelator is designed to transform the temporal noise into uncorrelated components with equal amplitude at all spatial frequencies (corresponding to "white noise"). In addition, it also acts on the signals inducing a property known as *delta-correlation*, which aids in the precise localisation of the position of the signal. Once the image data have been decorrelated they are passed to the cross-correlator which detects the signal by the mathematical process of cross-correlation with a series of reference signals.

9.2.1 The Decorrelator

It will be recalled, that the temporal noise has been modelled by a first order Markov random field (MRF-1). Decorrelation of an image which is approximated by an MRF-1 is based on the following principal: an MRF-1 has a spatial frequency spectrum which declines exponentially with increasing frequency and because decorrelation is achieved by causing the spectrum to be flat, the spatial frequency spectrum of the MRF-1 must be multiplied by a function which increases exponentially as the spatial frequency increases (Figure 9.1). The spatial domain (that is, the normal x, y coordinates of the image) equivalent of this operation is the mathematical process of convolution between the image function and a function known as a *spatial differential operator* (Faugeras and Pratt, 1980).

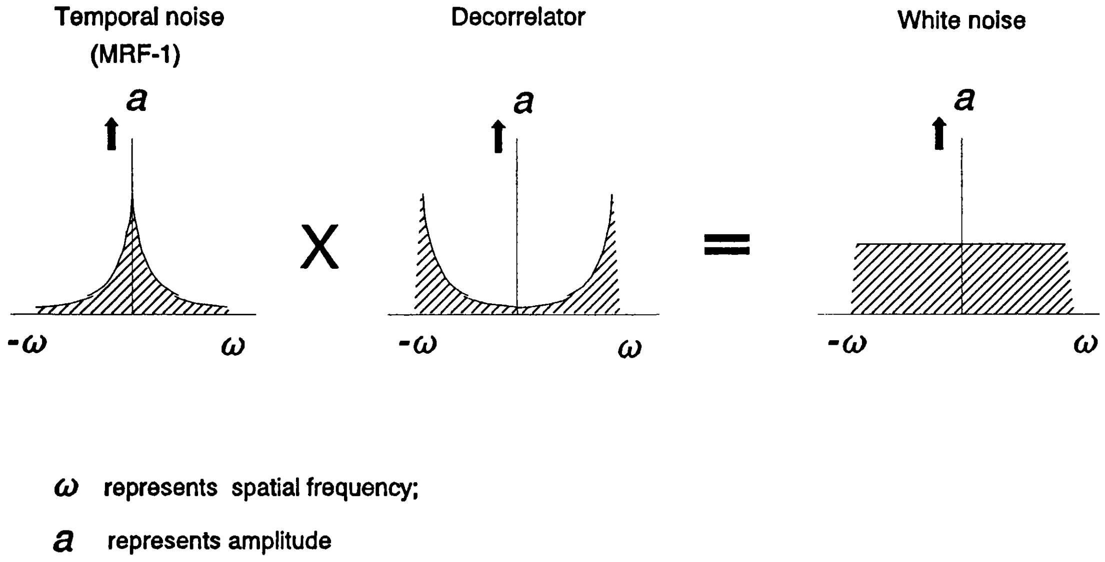


Figure 9.1 Two-dimensional representation of Decorrelation of the temporal noise in the Fourier domain.

A pilot investigation, conducted to determine the adequacy of decorrelation of the temporal noise using the standard spatial differential operators (Laplacian and Sobel gradient operators) indicated that while some degree of decorrelation occurred, neither operator rendered the temporal noise statistically white. The reason for this appeared to be that both standard operators work on the basis that the internal correlation of the MRF-1 is close to unity which is clearly not valid for the temporal noise.

It was necessary therefore to devise a specific spatial differential operator which takes into account the internal correlation of the temporal noise. The derivation of this operator is given in Appendix A.

To implement the decorrelation filter we must first know the magnitude of the internal correlation of the temporal noise, given by its one-step correlations, $\rho_{1\text{-step}}$, in the horizontal and vertical planes of the image. Unlike the form of the ACF, these cannot be deduced precisely from the theoretical temporal noise model. In addition, it is likely that the values of the one-step correlations will vary depending on the type of radiographic recording system and the spatial resolution or pixel size which is used to represent the image. Therefore an empirical approach was adopted to estimate the one-step correlations using samples of the temporal noise derived from accurately registered radiographs from the subjects in data set 1a. The description of the estimation procedure is given in Appendix D.

9.2.2 The cross-correlator

The cross-correlator is designed to process the decorrelated signals and noise and provide an output which is proportional to the probability of a true match occurring at each possible site of match.

The operation performed by the cross-correlator can be implemented either as a local correlation (or convolution) process in the spatial domain or as a global technique in the Fourier domain. However, because of the need to detect low contrast features, and to impose statistical stationarity on the images, the correlation process is more conveniently carried out in the spatial domain, thereby allowing for locally-varying correction of the mean and variance.

Many different correlation methods are available to perform detection processing of one- and two-dimensional signals (Jordan et al., 1989). However, because of possible dependence between the signal and temporal noise revealed by the validation study, a correlation method is required that is both statistically robust and relatively insensitive to signal dependent noise. These features are provided by the group of computational methods known as least-squares or sum of squared difference (SSD) correlation functions (Secilla, et al., 1988). The specific function chosen for this study was the *normalised SSD correlation algorithm* which allows statistical stationarity to be imposed on the images as an integral part of the cross-correlation process (Wiles and Foreshaw, 1993). The mathematical form of this function and its relationship to the correlation coefficient are given in Appendix E.

9.2 IMPLEMENTATION OF THE DETECTION SYSTEM

Foremost among the methods for implementing matched filter detection in the spatial domain is *template matching* (Chaudhuri et al., 1989). In this method the signal to be detected consists of a small area or *template* drawn from one of the images to be matched. The data in the template is then compared sequentially with all possible matching sites of the same size and shape in the second image. The correct match is determined by locating the point of maximum similarity between the template and an area of the same size in the second image. Similarity is measured by the numerical value of the cross-correlation function.

9.2.1 Template matching scheme

Before a template matching scheme can be implemented the following features must be determined:

(a) the size of the template;

(b) size of the search region in which the match will be sought;

(c) method of template selection;

(d) the way the search will be conducted.

9.2.2 Template size

The template size, in both physical terms and in terms of the number of pixels, was determined by a compromise between the need to ensure a large enough sample to minimise errors in the calculation of the correlation function while keeping the template as small as possible for ease of data handling and to limit the template to an area containing significant amount of stable structural detail.

The optimum size of the template is usually between 15 x 15 and 30 x 30 pixels dependent on the geometric distortion between the images (Mostavi and Smith, 1978b). However, to minimise the problems of data handling and to ensure a high density of reference sites per unit image area, the template was set at 15 x 15 pixels. While the shape of the template can be varied to adapt to the shape of the signal, a square template was adopted for computational convenience.

9.2.3 The search area

The size of the search area was set by the need to be certain that should a match exist it would be contained within the area of the search. Without external information about the sites of possible stable structure, it was considered reasonable to assume that the site of possible match could only be determined within ± 4 mm in any direction although, in certain instances, even this wide margin might be optimistic as indicated by the work of Baumrind et al. (1992). It was decided therefore to conduct the search in a region 12 mm square, centred on the anatomical site corresponding to the origin of the template data. For example, if the template was drawn from the region beneath the apex of the first molar in the later film, a possible match would be sought in a 12 mm square centred at the same position in the earlier film, as shown in Figure 9.2.

9.2.4 Template selection and search methods

Template matching, using complicated correlations functions, is a computationally intensive process requiring many millions of calculations (Jordan et al., 1989; Pratt, 1991). In addition, for any given template, the larger the search area the greater the probability of detecting a false peak which exceeds the height of the peak at the true site of match. For these reasons it was considered important to search for a match only for those templates with a high *a priori* probability of providing a true match and a low *a priori* probability of providing a false match.

Both probabilities depend on the output signal-to-noise ratio (SNR_{out}) of the cross-correlator which is directly related to the input signal-to-noise ratio (SNR_{in}) (Ryan et al.,

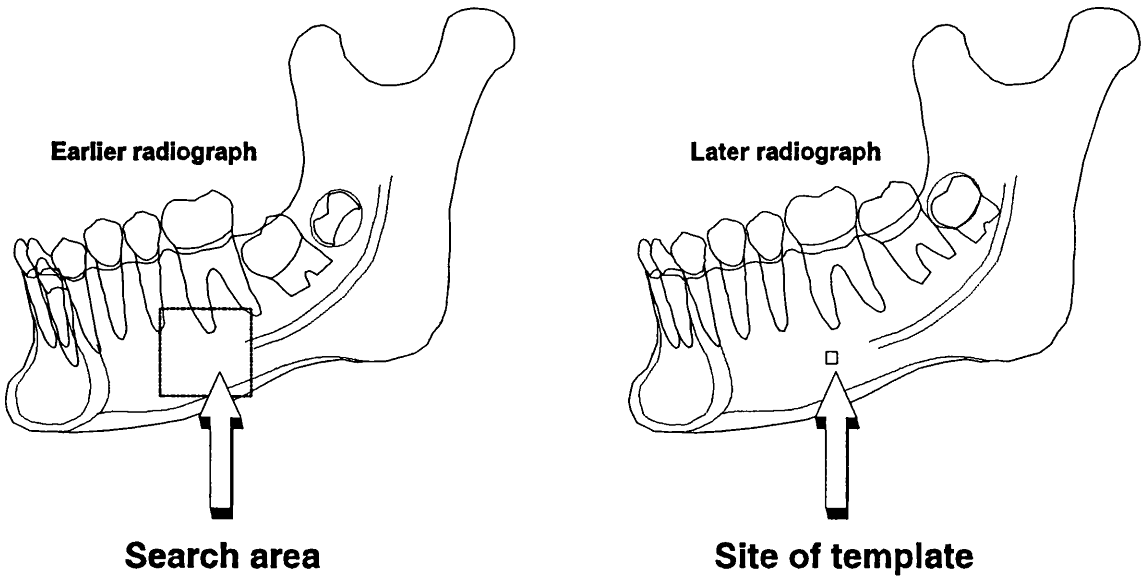


Figure 9.2 Diagram showing the relationship between template and search region

1991). Because the temporal noise is statistically stationary the SNR_{in} varies directly with the signal variance. Unfortunately, although decorrelation increases the image variance there is an increased probability of detecting a false match when decorrelated images are cross-correlated, especially where the SNR_{in} is initially low - that is, where the bone image differs in the two radiographs. Accordingly, to maximise the probability of true-detection and reduce the probability of false-positive detection, the system was implemented as a four stage procedure:

stage 1 - select templates from regions with highest SNR_{in};

stage 2 - maximise the SNR_{out} to detect the presence and approximate site of match without decorrelating the images;

stage 3 - decorrelate the template and search region;

stage 4 - accurately locate the site of match using the decorrelated images and searching a small area around the most probable site located in stage 2.

9.3 THE STAGES OF THE DETECTION SYSTEM

9.3.1 Stage 1: template selection

The image regions with the highest SNR_{in} and the greatest probability of being matched are those with the largest variance. Accordingly, the templates used by the system are located by processing the later image to detect the regions of highest variance and then selecting the templates from sites at the same regions in the original unprocessed image.

The variance detection method consists of applying a two-dimensional gradient operator to the later image to produce a gradient map which is then thresholded to reveal the regions of greatest local variance. Because it is important to avoid detecting regions with a high local variance due to noise rather than signal, the Sobel gradient operator (Kittler, 1983) was used followed by a median threshold to select the regions of greatest local variance. The Sobel gradient operator is a differential gradient operator which is convolved with the image to produce a new image whose pixels have a grey level that is proportional to the local variance or gradient across the surrounding eight pixels.

9.3.2 stage 2: high probability detection

The decorrelator produces images with inherently high variance which allows very precise localisation of the peak. Unfortunately, there is a greater probability of detecting a false-match with high variance data than with low variance data (Knapp and Carter, 1974). Thus, where a match is sought over an area of 12 mm x 12 mm at a pixel size of 0.13 mm x 0.13 mm, 5976 cross-correlation values will be obtained. Consequently, even if the probability of detecting a false peak which exceeds the height of the true peak is only one-in-a-thousand, on average a wrong peak will be located almost six times as often as the true peak.

The second stage of the detection system is designed to avoid this problem by searching for a match without decorrelating the images. Thus, the detection search in the second stage is conducted with lower variance images which, although they provide a low probability of detecting a false match, can only provide an approximate location for the match in the presence of noise.

To reduce the probability of falsely detecting a stable site during stage 2 of the detection system, a threshold was set for the acceptance of the peak as indicating a true match. The threshold was chosen such that only rarely does a peak value obtained from an incorrect match reach above this threshold, giving rise to a false match.

The threshold can be determined from theoretical considerations if the probability density functions of correct and incorrect outputs are known or assumed. However, in the absence of previous studies on detection of structures in bone the threshold was determined empirically. This was done by measuring the distribution of the correlations between the template and the regions surrounding the true peak and is reported in Appendix F. Using the results of that experiment, the thresholds were set at $\rho=0.68$ and $\rho=0.75$ for the mandible and maxilla respectively.

9.3.2 stage 3: decorrelation of the template and search region

The third stage of the detection system involves the decorrelation of the template and search region for which a match has been detected in the previous stage. The decorrelation is carried out by passing the images through the decorrelator. That is, by convolving the decorrelator impulse response array with each image in turn. However, this involves two problems.

The first problem is one of unfiltered noise. Despite the initial noise filtering provided as the radiographs were converted to digital signals, some residual noise will inevitably remain. Electric-line noise and quantum noise have largely high frequency components which are amplified during decorrelation because the decorrelation operation accentuates the highest frequency components of the images. The result is a differential amplification of noise which tends to overwhelm the image signals. To tackle this problem, the high frequency noise must be suppressed prior to decorrelation. This is achieved by convolving the data from the template and search region with a Gaussian impulse response array, which has the effect of replacing each pixel in the images with a Gaussian weighted sum of the eight surrounding pixels.

Secondly, as explained in section 3.4.3., spatial domain convolution leads to a band of pixels with spurious values at the border of the decorrelated (output) image, half the width of the impulse response array wide. To overcome this "border effect" the convolution operation is only computed when the impulse response array is contained fully within the confines of the input image. Thus, the output image will be smaller than the input image by the height and width of the impulse response array. Consequently, the convolution is conducted over an image area larger than the ultimate search region and template; from which the search region and template are then extracted leaving the spurious pixels behind.

9.3.4 stage 4: accurately locate the site of match

The fourth stage of the detection system is required to provide the accurate location of the matching structure in the earlier image. The accurate location is carried out using the decorrelated versions of the template and search region for which a match has been detected and approximately located in stage two. However, this search is conducted over a much smaller region around the location of the cross-correlation peak found in stage two. Because the true peak has been approximately located the size of search area can be reduced to an area just greater than the error of location in stage two. This error is described by the *correlation length* which is variously defined as the standard deviation of the width of the ACF (Mostavi and Smith, 1978b) or full-width at half-maximum height (FWHM) of the ACF (Bogler, 1986). For an exponentially declining ACF the second definition in more useful. However, in practice, the two values will be similar.

The FWHM correlation length for screen/film radiographs of bone was estimated using the data derived in section 8.2.2.4. The maximum FWHM of the ACF found was 7 pixels (mean and median 4 pixels). Accordingly, to be confident of locating the true peak the search in the third stage of the detection system was set at 15 pixels on either side of the peak detected in stage 2.

9.4 PRACTICAL IMPLEMENTATION OF THE DETECTION SYSTEM

The practical implementation of an image processing system is most conveniently carried out by computer.

In processing with a digital computer, the operations performed on the images may be carried out either by special purpose hardware consisting of digital logic circuits or alternatively by a general-purpose computer programmed in "software". The great flexibility of software makes this method the most appealing means of implementing a system and this approach has been adopted in this study.

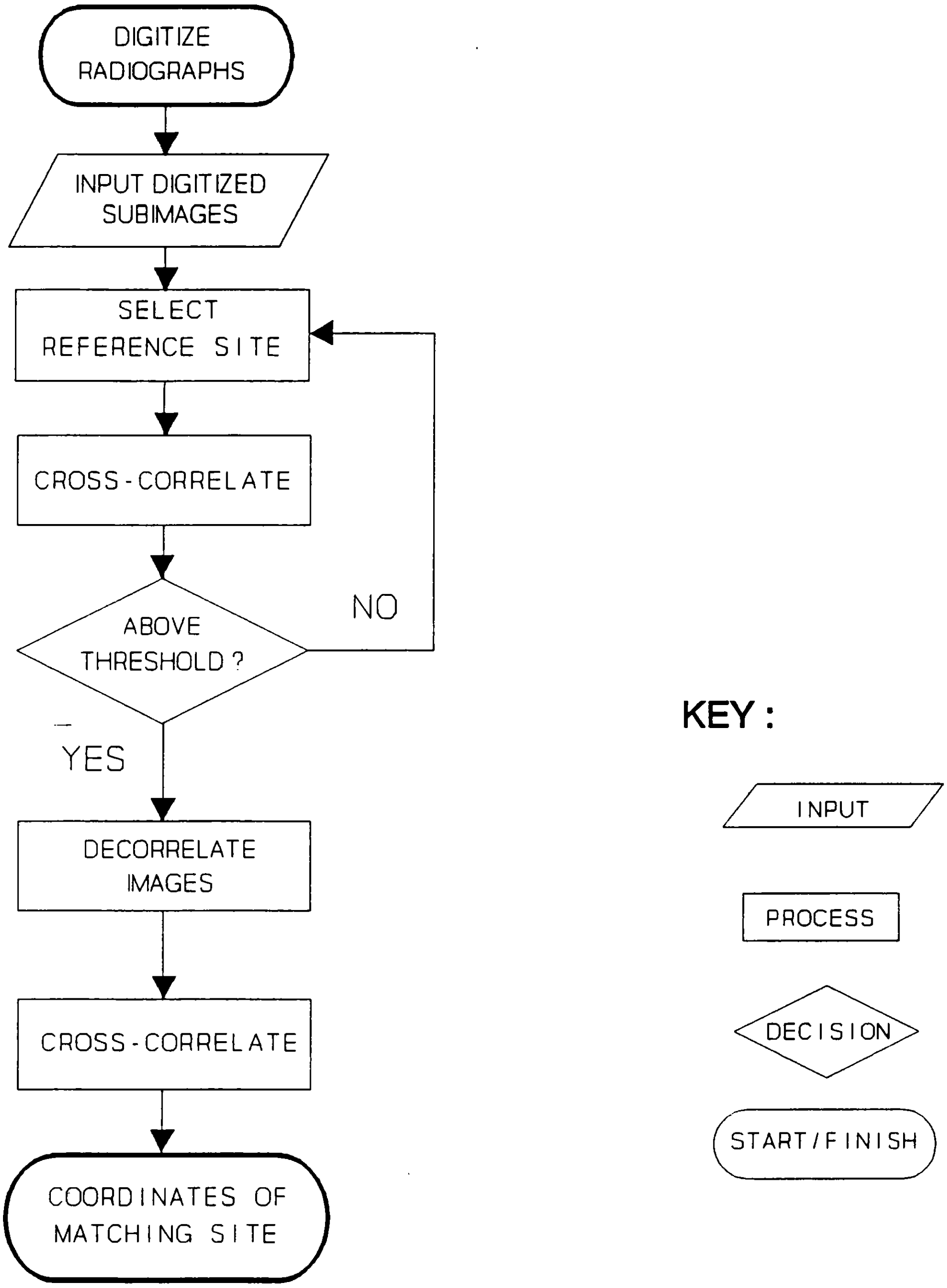


Figure 9.3 Flow chart showing the principal processing requirements as a sequence of operations.

9.4.1 System design

The design of a system to be implemented in software consists of two fundamental parts:

(a) detailed specification of the processing requirements;
(b) segmentation of the processing into programs.

9.4.1.1 *Processing requirements.-*

The processing requirements for the implementation are: input data, output data, and the four principal stages defined in section 9.3. These are shown in the form of a flow chart in Figure 9.3.

Requirements for template selection.- Digitized images from the later radiograph of each pair are convolved with a 3 x 3 Sobel gradient impulse response array. The resultant gradient image is then contrast-stretched (to ensure the gradients cover the full dynamic range of the image) and thresholded at the median grey level. Those pixels below the threshold are set to zero; those above the threshold to unity. The resulting binary image, which indicates the regions of greatest local variance, is then used to guide the selection of suitable templates from the original (unprocessed) later image. The results of each procedure are shown in Figure 9.4.

Requirements for reference site detection.- The template image and the much larger subimage in which the match is to be sought, are cross-correlated by the computer and resulting cross-correlation values stored in a correlation matrix. The matrix is then scanned by the computer to locate the position of the maximum value. The maximum value is then displayed together with a grey-level or colour coded two-dimensional representation of the correlation surface and a three-dimensional display of the correlation surface (Figure 9.5). The display of the entire correlation surface allows the peak value to be interpreted in relation to the size and position of the correlation peak or peaks. If the cross-correlation maximum fails to reach the threshold for acceptance, the procedure is repeated with the next template.

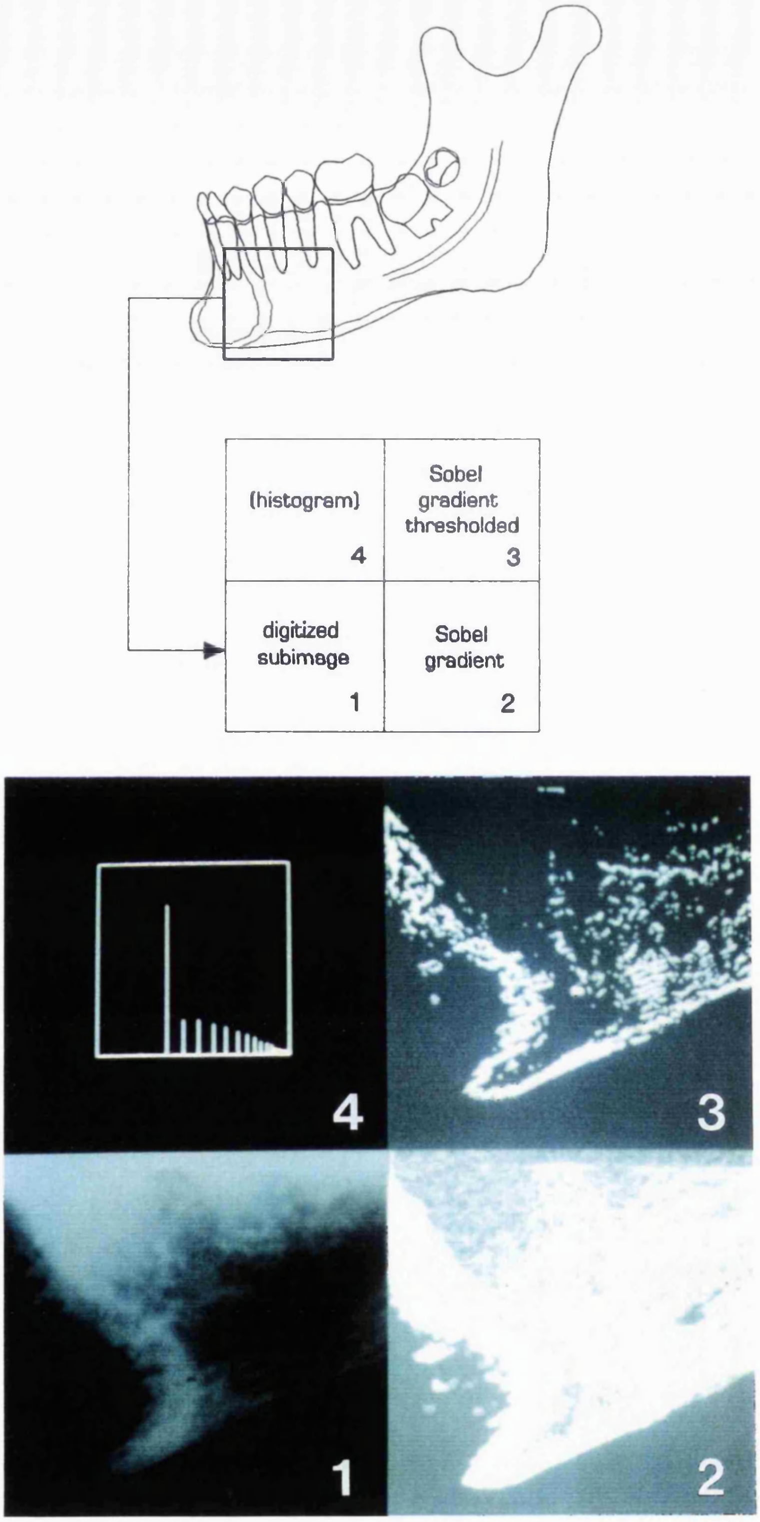


Figure 9.4 Sobel gradient processing of subimage to identify potential template regions.

Requirements for image decorrelation.- Where the peak value of the cross-correlation exceeds the detection threshold, the template and search subimage are decorrelated prior to passing the images on to the final stage to locate the precise position of match. The subimage from which the template was drawn and the search subimage are duplicated from the original image data held in the computer's memory. These two subimages are then convolved with a 3 x 3 Gaussian noise-cleaning matrix also held in the computer's memory. The pixel values in the resulting images are redistributed to occupy the full dynamic range of the grey scale ("contrast stretched"). These contrast stretched images are then convolved with a second 3 x 3 matrix held in the computer's memory. This matrix contains the nine integers which are the coefficients of the decorrelator impulse response array.

The output data (which cannot be displayed as images because they contain negative values) are multiplied by: one, at all sites corresponding to the final search region (30 x 30 pixels) and final template (15 x 15 pixels); and by zero, outside these regions, thereby removing the spurious border pixels. The remaining data are then compressed to occupy the normal eight-bit range of the image bitstack (that is, the computers image memory) and stored in memory. These decorrelated, reduced and compressed data are the new template and new search region.

Requirements for reference site location.- The new decorrelated template and subimage are cross-correlated by the computer. However, this cross-correlation is conducted over a 30 x 30 pixel region centred on the coordinates of the peak found in the previous cross-correlation. Again the resulting cross-correlation values are stored in a correlation matrix which is then scanned by the computer to locate the position of the maximum value. The precise coordinates of the maximum are displayed together with a three-dimensional representation of the new correlation surface (Figure 9.6). This display can be tilted and rotated to allow it to be viewed from any angle again allowing the peak value[†] to be interpreted in relation to the size and position of other correlation peaks.

[†] Because of the nature of SSD correlation, the "peak value" is infact a *minima* but is automatically displayed as a peak or *maxima* by inverting the correlation surface.

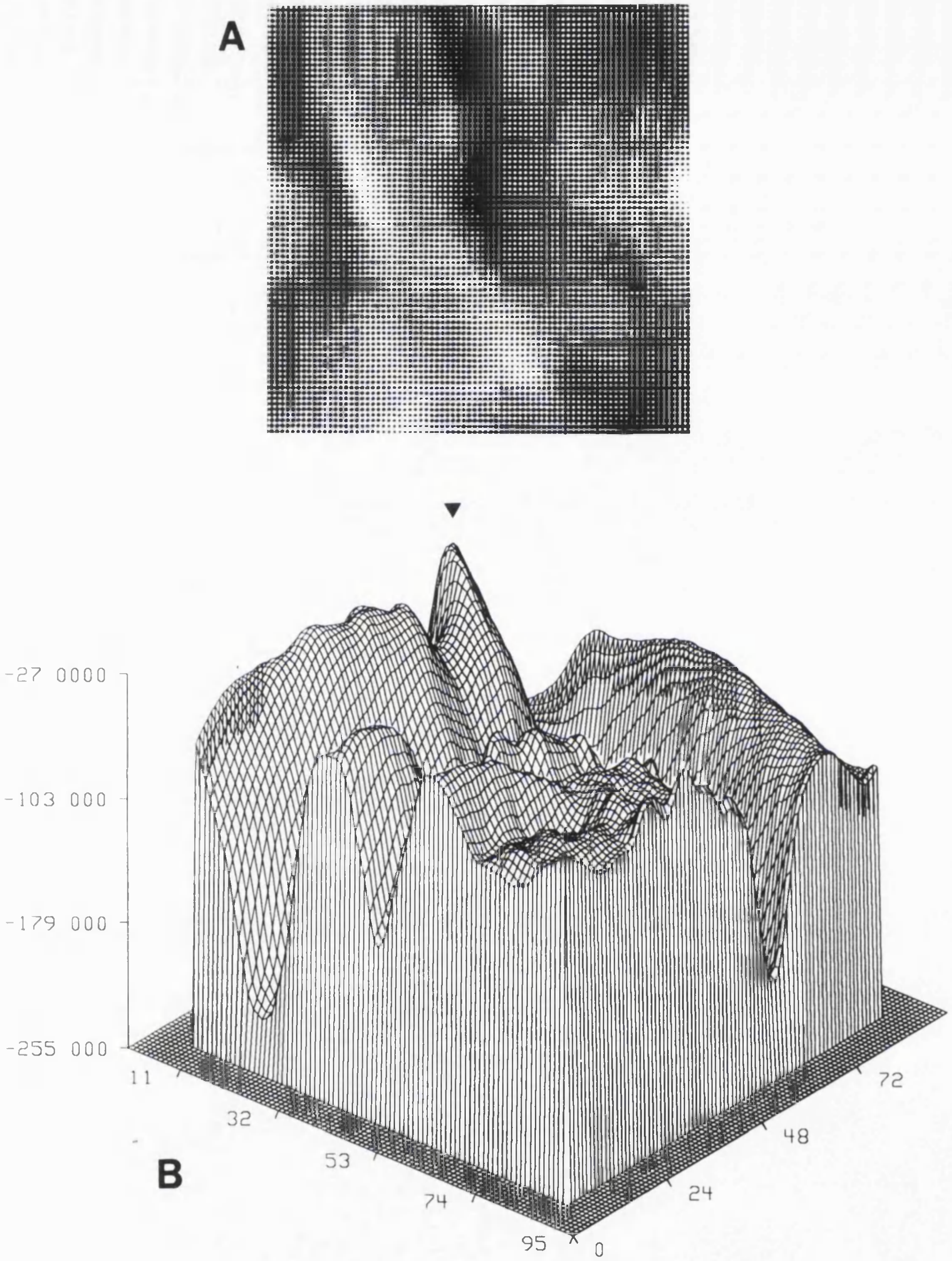


Figure 9.5 Output display from reference site detection stage: A) grey-scale coded 2-dimensional image of a correlation surface; B) 3-dimensional correlation surface.

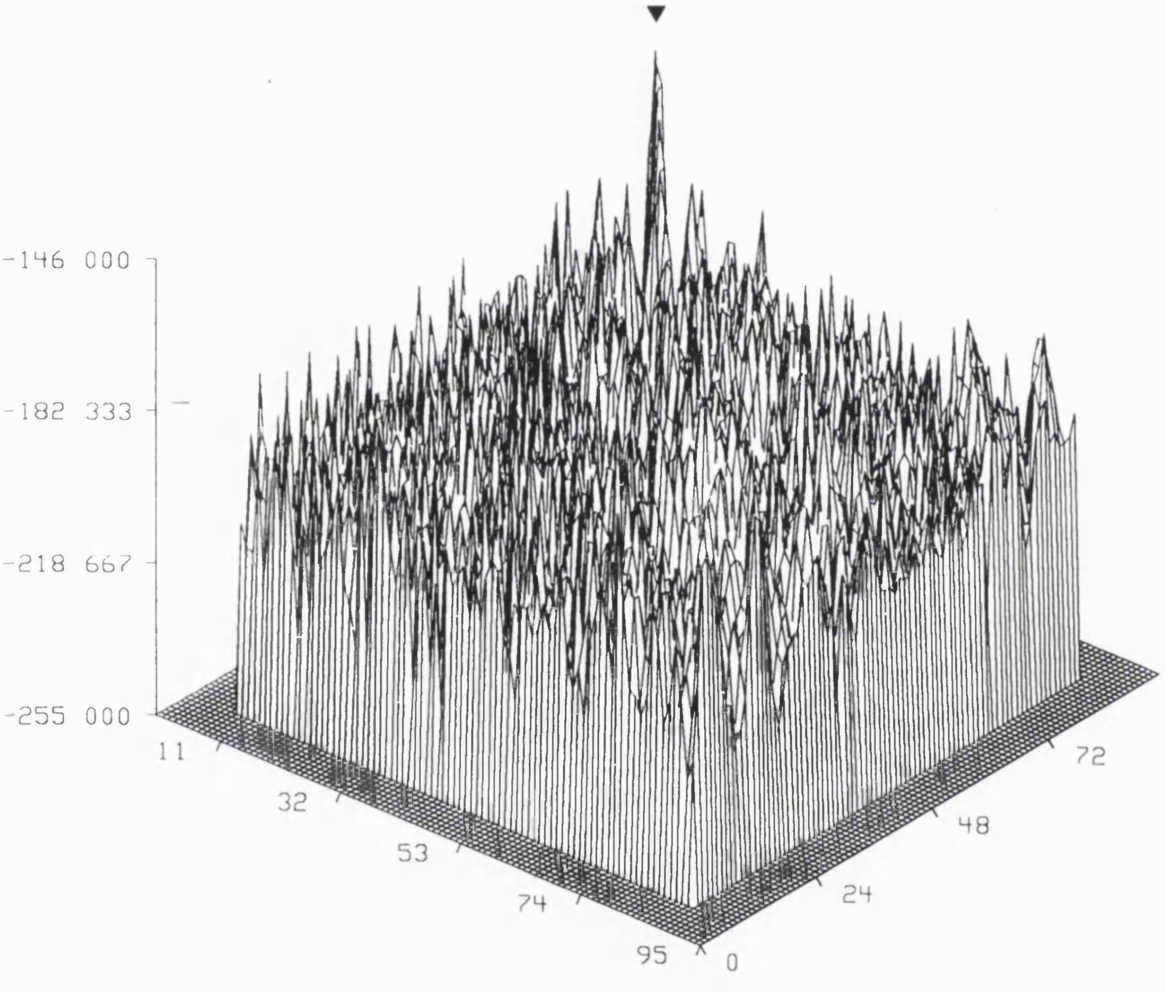


Figure 9.6 Output display from reference site location stage showing the 3-dimensional image of a typical correlation surface for decorrelated images.

9.4.1.2 *The segmentation and software programs*

The detection system was segmented into the five principal stages shown in Figure 3 and a software program was written to perform each stage:

image capture and noise filtering program - **GRAB**;
template locating program - **SELECT**;
image decorrelation program - **DECORR**;
correlation detection program - **CORR**;
correlation location program - **SIMPLE_CORREL**

The computer programs each went through a period of testing and refinement prior to inclusion in the prototype system. The testing was primarily conducted on the two sets of prototyping radiographs removed from data set 1 prior to the division of the data set into data sets 1a and 1b.

The source codes for the programs **GRAB**, **SELECT**, and **DECORR** were written in the high-level computer language YAM by the author. **CORR** and **SIMPLE_CORREL** were written in the computer language "C" by two experienced programmers (Dr C. Wiles and Dr D. Greenwood). **CORR** is a rapid execution program which calculates the cross-correlation without detailed estimation of the exact site of match. **SIMPLE_CORREL** is a more detailed correlation program which allows accurate (subpixel) estimation of the position of the cross-correlation peak. The source codes for the programs can be found in Appendix G.

Once initiated, each program executes automatically, however, the programs are not interlinked. Thus, it is necessary to load each program at the appropriate point and to transfer (and where necessary, re-format) the images under interactive control.
While this slows down the process the calculation of the cross-correlation, involving hundreds of millions of computer operations, is the most time consuming and limiting part of the process.

9.4.2 The prototype detection system

The detection system was implemented using the computer-based image processing system described in chapter 7. The images were acquired using the equipment previously described and at the same spatial and grey-scale resolutions.

9.4.2.1 *The basis of the practical procedure*

The general principals of the practical procedure are as follows. The radiographs are sorted into a time-ordered sequence from earliest to latest and numbered in sequence (1,2 ....n). One radiograph is selected as the base line reference (usually the earliest). All other radiographs are then paired in turn with the base-line reference and processed to locate reference points common to each pair of images. Thus, the detection system operates to detect reference points for pairs of radiographs.

9.4.2.2 *The system algorithm*

The operations performed on the images consist of a series of mathematical computations specified by the software programs. The sequence of operations needed to implement the detection system using the software programs is called an *algorithm*. The algorithm proceeds in 19 key stages:

Start:

1. Place radiographs in turn beneath camera. Each pair of radiographs should be orientated as closely as possible to the expected final registration position.
2. Run program **GRAB**
3. Once satisfactory images stored exit **GRAB**
4. Duplicate images using YAMTOOL function
5. Run program **SELECT**
6. select a template and exit **SELECT**
7. To UNIX shell edit window, convert image one and two from YAM to RAW format.
8. Run correlation program **CORR**.
9. Display (unwhitened) later image in the YAM window, locate centre of template (region to be detected) and call up YAM cursor function. Place cursor at centre of proposed template and enter coordinates at the **CORR** prompt in UNIX shell window.
10. Read out central coordinates of site.
11. Check extent of proposed window (15 x 15 pixels).
12. Enter coordinates of template and search position together with filename for the correlation matrix (for example *cormat.001*).

13 Display output correlation matrix using *seeplot* diplay function of PLOT3D(13) program. Verify presence of regions above threshold; If no region above threshold repeat from stage 2 with new site.
14 If a region exists above the threshold, store: coordinates of peak; and correlation matrix. Exit **CORR**.
15 Run **DECORR** using duplicated images from stage 3.
16 Exit **DECORR**.
17 In UNIX shell edit window, convert decorrelated images to pixrect format (.pr).
18 Run correlation program **SIMPLE_CORREL**, enter same coordinates as before, template size (15 x 15) and subpixel dimensions of the search, at the prompt
19 Store: coordinates of peak; and correlation matrix. Exit **SIMPLE_CORREL**.

Repeat sequence 1-19 for next site.

When sufficient templates have been matched (typically twenty for each 1.55 cm^2 subimage) the offset coordinates of the templates are averaged to provide mean offset coordinates for each pair of subimages. The mean offset coordinates are used to register the original digital images held in the computer's memory. This is carried out by simply instructing the computer to translate the subimage from the later radiograph by the offset coordinates using the (interactive) YAM shift function. While this procedure should provide an accurate alignment of each pair of subimages the mean coordinates may differ between pairs of subimages because of the rotation, and other higher order distortions between the original images.

9.4.3 Comments of the detection system

It is important to note that, unlike (digital) subtraction methods (Jeffcoat et al., 1984; Gröndahl, 1987), the detection system described here does not require the careful standardisation of radiographic exposure or processing techniques to produce comparable contrast in the serial radiographs. This is because of two factors. First, the local contrast and brightness are forced to correspond in the two images by the normalisation process which occurs as part of the SSD cross-correlation. Secondly, the detection of matching structural details is based on the comparison of spatial gradients or grey-level boundaries and not on correspondence between the image densities.

The prototype detection system has been organised for the registration lateral cephalometric radiographs which allows large (1.55 cm^2) subimages to be aligned by translation alone. This is because of the precise control exerted over the projection geometry by the use of a cephalostat and large source to object distance (Jeffcoat et al., 1987). However, for other radiographic views the size of subimages will need to be reduced to ensure that registration can be achieved by translation of one image relative to the other.

9.4.4 Limitations of the detection system

The detection system as described here is rudimentary and has been constructed simply to implement the theoretical system developed in chapters 5 to 8. As such it represents a prototype system for which many improvements are possible. There are however, some important limitations which should be mentioned.

The detection system as presently organised has no facility to correct rotational misalignment between the images (rotation in the plane of the film); nor does it have the capacity to register images which are distorted by changes in the orientation of the bone to the x-ray source or film. Computer programs are available to perform this type of "elastic matching" (Bajcsy and Kovačič, 1989) but they would only be of use for images of bone with limited depth.

In addition, the detection system has no facility to detect and eliminate "bad" matches. Although gross blunders will generally be obvious and can be rejected, under certain circumstances (for example, rotational misalignment between the images) the disparity between the offset coordinates of different templates may make it difficult to separate blunders from "good" matches.

At present the system has been implemented only on a Sun 3/160 computer with a default allocation (under YAMTOOL) of 300 megabytes of memory for image storage. With a few modifications to the computer programs, it is likely that the detection system could be implemented on a standard personal computer (for example, an IBM PC compatible personal computer equipped with a 80486 central processing unit(14)) with little loss of speed.

CHAPTER
10

TESTING THE DETECTION SYSTEM: PERFORMANCE EVALUATION AND ERROR ANALYSIS

10.1 INTRODUCTION

At its output, a matched filter detection system provides a set of statistics. Each statistic represents a measure of the similarity of the signal (a subsection of the first image) to a possible matching site in the second image. The position of the peak output indicates the point at which the probability of a true match is greatest. However, the point of maximum probability may not be the true position of the corresponding structure. The final decision has to be made to accept or reject the position of the peak as the true site. This is therefore, a statistical problem: consequently, even if the processor performs perfectly according to its design, the detection of the presence and position of true match is always approached with some degree of uncertainty.

The probability of error is dependent upon several factors the most important of which is the level of the noise. This noise leads to statistical fluctuations in the measured correlation function which produce two distinct types of error in detecting and locating the peak of the correlation function (Figure 10.1).

The first type of error occurs when an incorrect lobe of the correlation function is selected (that is, the wrong "peak" is selected). This is referred to as a "false

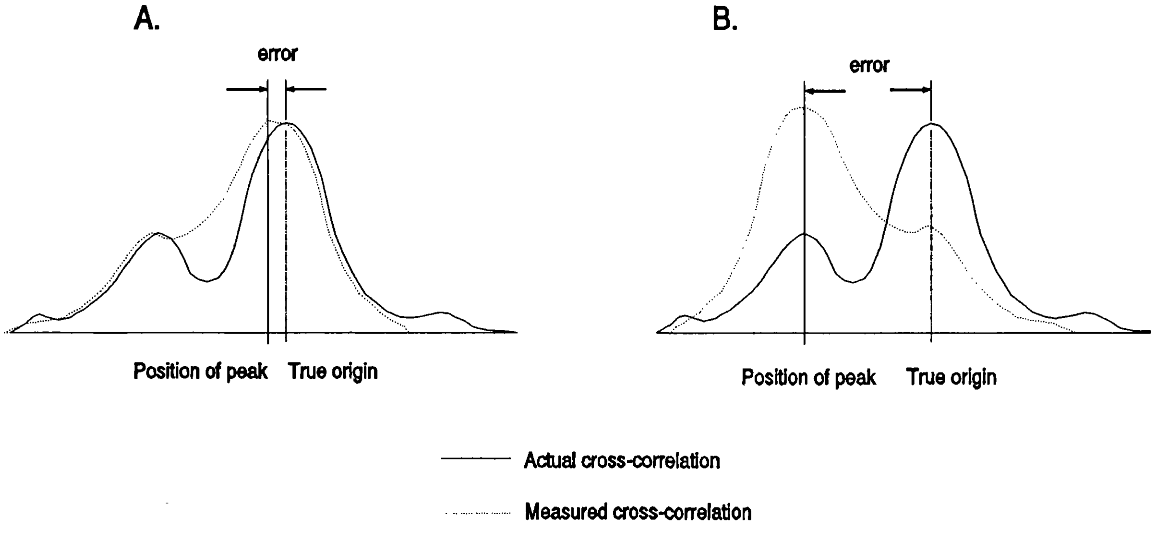


Figure 10.1 Diagrammatic representation of the two types of correlation error: A) local registration error; B) false acquisition error.

acquisition error". The other type of error arises when the correct "lobe" of the correlation function is selected but the peak of the lobe does not coincide with the correct position. This type of error is referred to as a "local registration error".

The two types of error tend to occur in different circumstances. In general, local registration errors occur where the correlation function is broad, while false acquisition errors occur where the correlation function is narrow or sharply peaked. Shallow peaks are less susceptible to false acquisition errors but give poor precision of location in presence of noise.

10.1.1 Sources of the noise

The noise which leads to detection and location errors arises from several sources in addition to the additive noise accompanying the signal. The various sources of noise in correlation-based detection systems may be conveniently categorised into one of the four classical sources of error in systems based on theoretical models (Von Neumann and Goldstine, 1947). These are:

(a) Modelling assumptions;

(b) Input data to the system;

(d) Rounding of arithmetic calculations;

(c) Approximations of incomputable processes or functions, by computable processes or functions.

10.1.2 Modelling assumptions

These have been explored in chapters 5 and 6, and their validity examined in chapter 8. Where the modelling assumptions did not correspond to real data it has been possible to mitigate this by transforming the data to match the assumptions of the model. However, the basic premise underlying the model is that detection takes place in a two-dimensional system. This is clearly not the case because although the radiographs are two-dimensional they represent three-dimensional objects. Unfortunately, there is no unique way to collapse the three-dimensions of the object into the two-dimensions of the radiographic image. To allow the models to be tractable the underlying assumption is made that the serial radiographs are spatially congruent that is, that they fit into each other in two-dimensions. Because of the difficulty of maintaining the same recording geometry at each exposure, this assumption is always violated to some extent. The sensitivity of the detection system to violation of this basic assumption will be examined in the experimental section using a physical simulation.

10.1.3 Input data

Noise which contaminates the input data is often propagated as the data are carried through the system leading to errors in the output - in this case, errors in the detection and location of stable structure. The noise at the input arises from the general noise in the radiographs and the additional electronic noise induced during the conversion of the images to a digital format. The reduction and control of this noise has been discussed in chapters 7 and 8. However, although it may be overcome to a limited extent by averaging of the input data and applying noise reducing filters, the effects of this type of noise are most pronounced for digital images consisting of small samples, as is the case in the present study. Consequently, a compromise must be established between the complexity of using larger samples and the level of noise which is tolerable. Such attempts have been made in chapters 7 and 9.

10.1.4 Rounding of arithmetic calculations

Rounding and truncation of arithmetic calculations can be significant sources of error in correlators with limited precision and form an important part of the errors resulting from the direct operation of the computer-based detection system.

Rounding and the effects of finite numerical precision are most serious in the calculation of the cross-correlation because they add white noise to the signals. In addition, if the "difference equations", which form the basis of a digital processor implemented by computer, are implemented with finite numerical precision the resultant system will be non-linear, which leads to a serious theoretical problem (Oppenheim and Shafer, 1989).

These sources of noise are assumed to be minimal in modern computers where individual values are stored to the first 15 correct digits. However, "numerical instability" can result if simple computer languages, which do not employ double-precision arithmetic, are used to write the programs on which the system is based (Digby and Kempton, 1989). This is not the case in this study where all programs have been written in languages which employ high precision arithmetic for all calculations.

10.1.5 Approximations of incomputable processes

For complex systems to be workable at all, it is often necessary to deal with incomputable processes or mathematical functions. Approximate but workable solutions are therefore used instead. The noise originating from this source is usually considered extremely difficult to remove and with models of significant complexity it is difficult to avoid completely (Meyer, 1984). Methods of minimising errors caused by this noise form the basis of the branch of applied mathematics known as numerical analysis.

The principal approximations in the present detection system are:

(i) Approximation of the decorrelator impulse response by 3 x 3 array with integer values;

(ii) Approximations in the calculation of the cross-correlation.

10.1.5.1 *Approximation of the impulse response*

The impulse response of the decorrelator developed in this study is the smallest that it is practicable to use. Its effect extends to the 8 pixels around the point being processed and each of its nine constituent coefficients consists of integer values. These limitations and approximations are imposed by the computer software (YAMTOOL) and can only be improved upon by using alternative software programs. However, the effect of this noise is only of practical importance if the decorrelator fails to decorrelate the temporal noise. This was tested in a series of pilot studies prior to implementing the detection system. In all cases, the temporal noise was rendered statistically white, that is, the $\rho_{1\text{-step}}$ of the temporal noise was not significantly different from zero at the 5% level ($p > 0.05$) for the 15 x 15 pixel samples examined.

10.1.5.2 *Approximations in the calculation of the cross-correlation*

These approximations result from the small sample size employed to calculate the cross-correlation values and are unavoidable because the "true" cross-correlation is based on an infinite sample. The result of computing cross-correlation functions with finite samples from a random signal field is the production of a specific type of noise known as *correlator self-noise* or *intrinsic noise* (Ryan and Hunt, 1981; Barrett and Swindel, 1982b). Although it is termed "noise" it is in fact a deterministic effect which is unaltered by filtering or cleaning the data.

The two methods of reducing correlator self-noise are: to increase the sample size; and to increase the range of spatial frequencies included in the data sample. There are obvious practical limits on increasing sample size and the spatial frequencies in the sample are limited by the spatial resolution of the original radiographs. Accordingly, approximations are still required in any practical system designed to detect structure in clinical radiographs.

10.2 THE EVALUATION OF DETECTION SYSTEMS

During the operation of any detection system, at each location where a match is sought, one of four situations may occur (Swets and Pickett, 1982):

(i) The true match is detected (true-positive detection);

(ii) The true match is missed (false-negative detection);

(iii) A wrong position is classified as a true match (false-positive detection);

(iv) A wrong position is classified as a false match (true-negative detection).

An ideal detection method would only make decisions (i) and (iv). Therefore, the performance of a real detection method can be described in terms of the probabilities of the different outcomes. However, only two of these outcomes are independent, thus only two probabilities need be reported to provide a complete description of the detection performance - probability of detecting a true match and the probability of incorrectly classifying a wrong position as a true match (situations (i) and (iii)).

Classically, the relationships between these two probabilities is indicated by a *receiver operating characteristic* (ROC) curve as shown in Figure 10.2.

The ROC curve conveniently depicts the two outcomes over the whole range of probabilities independent of the underlying probability distributions (Swets and Pickett, 1982). An almost ideal detection method will produce a curve which lies close to the left-hand and top margins of diagram, only deviating from zero false-match probability where the true-match probability is close to unity (curve C Figure 10.2). For a detection or classification method which is no better than chance the curve will lie along the diagonal (curve A Figure 10.2). A practical system will produce an ROC curve between the two extremes of ideal detection and pure chance, as indicated by curve B Figure 10.2.

10.2.1 Evaluation of correlation-based detection systems

The performance of correlation-based detection systems is usually judged on two types of performance measure related to the two types of detection error. The first measure is the probability of locating the correct lobe of the correlation function - the *detection accuracy*. The second measure is the accuracy with which the offset of the

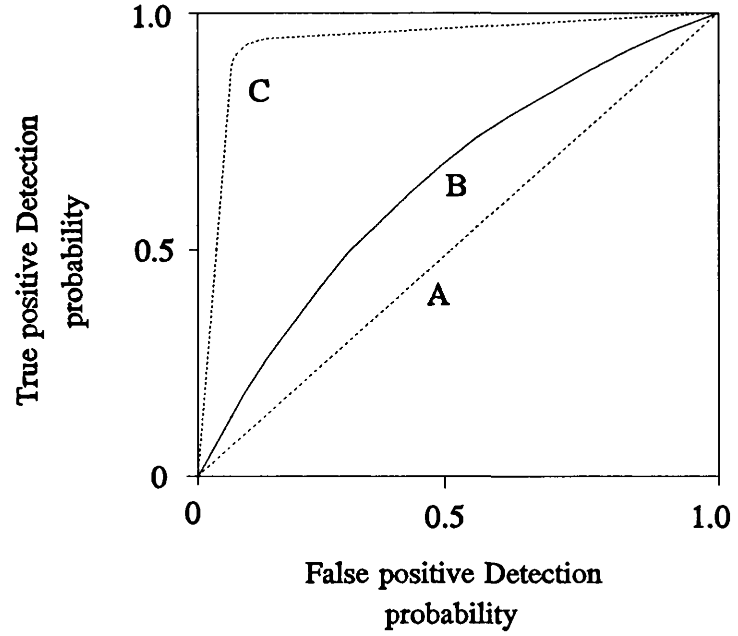


Figure 10.2 Receiver operating characteristic (ROC) curves: A) detection no greater than chance; B) practical detection system; C) ideal detection system.

peak of the correlation function corresponds to the true position - the *local accuracy*.

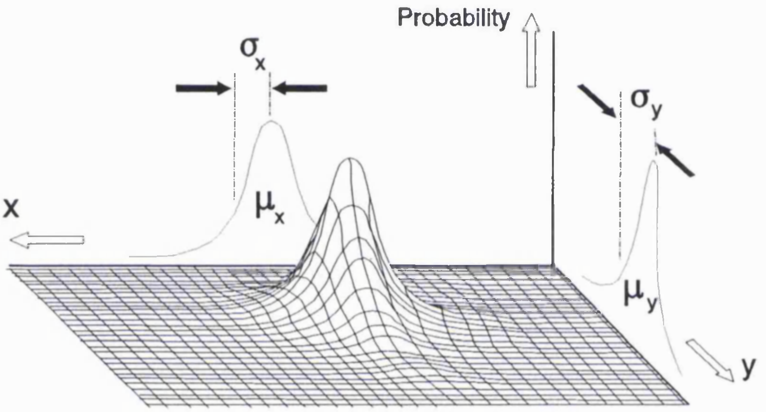
10.2.2 Detection accuracy

For any detection system two aspects of detection accuracy must be distinguished: first, the inherent discrimination of the system; and secondly, the accuracy obtained using a specific threshold (Swets and Pickett, 1982).

The former is determined by the probability of true-positive detection versus the probability of false-positive detection as the threshold value changes. The latter is measured by the true- and false-positive probabilities at the particular threshold used by the system. This second measure is referred to as the *accuracy at the operating point*.

10.2.3 Local accuracy

The accuracy with which the offset of the peak of the correlation function corresponds to the true position is described by two measures: the covariance matrix[1] of the local registration error; and the mean vector of the local registration error (Mostavi and Smith, 1978b). The covariance matrix describes the random component of the error and the mean vector describes the systematic error or registration bias. These are shown diagrammatically in Figure 10.3. However, where the random errors in the two planes of space are known to be uncorrelated these errors may be expressed more simply by the one-dimensional error variance in each coordinate axis.



$$\text{mean vector} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} \qquad \text{covariance matrix} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{pmatrix}$$

Figure 10.3 Diagrammatic representation of the parameters of a two-dimensional error distribution: μ, σ^2 and σ indicate mean, variance and covariance respectively.

---

[1] The covariance matrix is the two-dimensional equivalent of the variance of a one-dimensional distribution.

10.2.4 Methods of performance evaluation

Two types of approach to performance evaluation are possible: theoretical; and practical. Theoretical measures calculate the performance from theoretical considerations - that is the models, while practical measures determine the performance empirically. Practical measures are generally preferred to theoretical ones for three reasons.

First, the estimates derived from the models are usually overoptimistic, Secondly, models only represent approximations to reality and unknown sources of error will not have been included in the models. Finally, the analysis of complex models can make the prediction of error very difficult. To allow the theoretical analysis to be tractable they often require assumptions which are not compatible with the empirical data (Mostavi and Smith, 1978a).

10.3 EXPERIMENTAL EVALUATION OF THE DETECTION SYSTEM

10.3.1 Aims and objectives

The performance of the detection system was examined in four experimental investigations. Each investigation was undertaken to determine a different aspect of system performance:

(i) Detection accuracy;
(ii) Local accuracy;
(iii) Detection accuracy at the system operating point;
(iv) Effect of (affine) image distortion on detection accuracy.

10.3.2 Experimental measurement of detection and local accuracy

The detection accuracy of the system was estimated from a set of test images for which the true site of match was known by reference to stable tantalum markers. The proportion of true-positive, true-negative and false-positive templates classified by the detection system were used to provide statistical estimates of the true-positive detection probability (*sensitivity*) and the true-negative probability (*specificity*). The difference between the true position (determined from the tantalum markers) and that found by the system for positive matches was used to provide statistical estimates of the local accuracy of the system.

10.3.2.1 *Design of the experiments*

The performance of the detection system was examined separately for each data set (1a and the combined set [1b + 2]). This was done to maintain the independence of the data sets and thereby provide realistic bounds to the expected accuracy.

The reasoning behind this approach is that data set 1a has been used during the development of the system and consequently, should provide an optimistic estimate of the performance. However, testing the system on the combined data set [1b+2] should provide a pessimistic estimate of performance because none of these data have been used to design or refine the system. The true performance of the system is assumed to be bounded by these two evaluations (Fukunaga, 1990).

Unlike the experimental validation of the models, the data from the individual subjects were combined to provide an overall estimate for each data set.

10.3.3 **Materials and methods**

The test images were derived from pairs of serial lateral cephalometric radiographs of the mandible and maxillae obtained from the fifty-two subjects from data sets 1a, 1b and 2. Data sets 1b and 2 were combined to provide a single data set comprising subjects whose records had not been used during the development of the detection system.

10.3.3.1 *Image capture and preliminary processing*

The image capture and subsequent selection of reference sites was conducted in the manner proposed for the normal use of the system, that is without knowledge of the true registration position. However, the images of the tantalum implant markers provided a potential source of bias in the alignment of the radiographs during image capture. Accordingly, to prevent the implants being used to judge the true registration position, the implant images in each later radiograph were obscured with aluminium foil discs 8 mm in diameter, placed arbitrarily over each implant image and secured with adhesive tape.

The earlier radiograph of each pair was secured to the variable intensity light box beneath the video camera, with the image of the mandible in the field of view. The image was then captured and stored. The later radiograph was then aligned on the earlier radiograph using only the anatomical detail visible in the images. The earlier film was then removed and the later radiograph secured in place on the light box. The aluminium discs were then carefully removed while ensuring that the radiograph did not alter its position. The image was then captured and stored. This procedure was repeated for the maxillary images.

As in the validation study, images from the earlier and later radiographs of each pair were recorded and stored as a series of subimages (approximately 1.55 cm^2) of 96 x 96 pixels with a pixel size of 0.13 mm x 0.13 mm from pre-determined sites. For each subject two subimages were selected; one from each jaw. The choice of subimage was made randomly from the five anatomical sites in the mandible and two anatomical sites in the maxilla as described in section 7.5.7 (see Figure 7.5). The subimages from the later radiograph were duplicated by the computer and displayed with the corresponding subimages from the earlier radiograph. These subimages were then registered using the tantalum markers as described in section 8.4.1. The registration coordinates for each pair of subimages (one from the earlier, and one from the later radiograph) were then stored.

The subimages from the later radiograph were processed to locate potential templates as described in chapter 9. Five 15 x 15 pixel templates were then selected from each subimage. The position of each template was determined arbitrarily from the potential sites in the central 50 x 50 pixels of each subimage, thereby providing a total of 5 x 24 templates for each jaw in data set 1a and 5 x 28 for each jaw in data set [1b+2].

10.3.3.2 *Detection accuracy of the system*

The detection system was used to search for a match for one template for each subject in five separate trials. At each trial a match was sought in two search regions: one region known to contain the corresponding section of bone, and another region which did not contain the corresponding section of bone. The plan of the experimental method is shown diagrammatically in Figure 10.4.

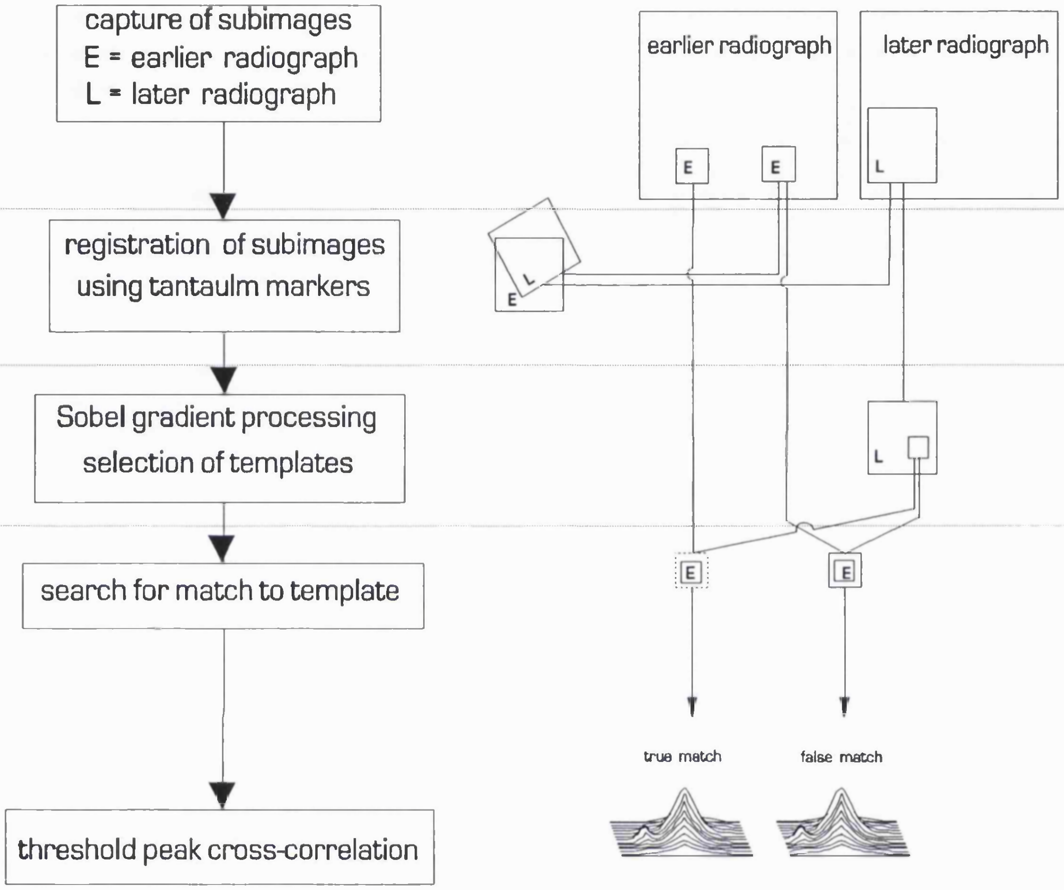


Figure 10.4 Flow chart and plan of experiment to evaluate the detection accuracy of the system.

Both search regions comprised an area of 25 x 25 pixels from the earlier radiograph. The first region was centred at the coordinates corresponding to the centre of the template as determined by the tantalum implants. The second search region was located in the same subimage as the first but at a minimum distance of 15 pixels from the coordinates corresponding to the centre of the template.

The detection search was conducted using only the first two stages of the system, that is, without decorrelating the images. Following each trial the resulting data were subjected to eight different thresholds for acceptance of the match. At each threshold detection was deemed to have occurred if the highest value of the cross-correlation exceeded the threshold. The eight thresholds used are shown in Table 10.1.

Statistical analysis.- The raw data from the detection accuracy trials were tabulated and used to construct individual ROC curves depicting system performance for the mandible and maxilla at each trial. The areas under the individual ROC curves were calculated from the tabulated raw data using the ARCUS data analysis program(12). The area under the ROC curves for mandible and maxilla from each data set were calculated by averaging the individual ROC curve areas from the five trials. Statistical analysis was performed using a non-parametric ANOVA with multiple comparison of means (Kruskal-Wallis test) to determine if significant differences existed between the data sets or for mandible and maxilla within each data set. In addition, the difference in detection accuracy from that expected by chance (ROC curve area = 0.5) was tested for significance using Wilcoxon's matched pairs signed ranks test.

Table 10.1 The thresholds used for the system detection accuracy experiments.

Threshold number	1	2	3	4	5	6	7	8
Mandibular Threshold (ρ)	0.15	0.30	0.45	0.55	0.60	0.68 †	0.80	0.90
Maxillary Threshold (ρ)	0.15	0.30	0.45	0.55	0.60	0.75 †	0.80	0.90

Key: † indicates the threshold at the system operating point.

10.3.3.3 *Detection accuracy at the operating point*

The results of the experiment just described should provide the evidence necessary to completely determine the detection accuracy at the chosen thresholds. However, the system operates not just by selecting those regions whose cross-correlation with the template exceeds the threshold, but by selecting the region with the maximum cross-correlation above the threshold. Consequently, to provide a more realistic evaluation of the detection accuracy at the chosen operating points a second experiment was performed where detection was attempted over a large search region and where the criterion for detection was that the maximum cross-correlation should occur within a small specified range of the corresponding position determined by reference to the tantalum implants.

The detection system was used to detect a match for one template in a region known to contain the corresponding section of bone. The templates used in this experiment were chosen at random from the five templates selected from each subimage (above). The search region comprised the entire corresponding subimage (96 x 96 pixels) from the earlier radiograph. Consequently, the search region was centred on the region corresponding to the 50 x 50 pixel area from which the template data were drawn.

As before, the detection search was conducted using only the first two stages of the system. In all cases detection was deemed to have occurred if the highest value of the cross-correlation exceeded the detection threshold at the operating point ($\rho = 0.68$ or $\rho = 0.75$ for mandible and maxilla respectively) and if its location was within 5 pixels of the position determined by reference to the implants. This procedure was repeated for both jaws in all subjects. The plan of the experimental method is shown diagrammatically in Figure 10.5.

Statistical analysis.- The data from this experiment were used to provide maximum likelihood estimates of the following, together with their 95% confidence intervals:

(a) probability of true-positive detection;

(b) probability of false-positive detection;

The 95% confidence intervals were determined from the values given by Highlyman (1962) for a binomially distributed error between two-alternative forced choices.

The true-positive / false-positive ratios recorded for the two data sets were tested for statistical significance using a χ^2 analysis to explore the possible superiority of

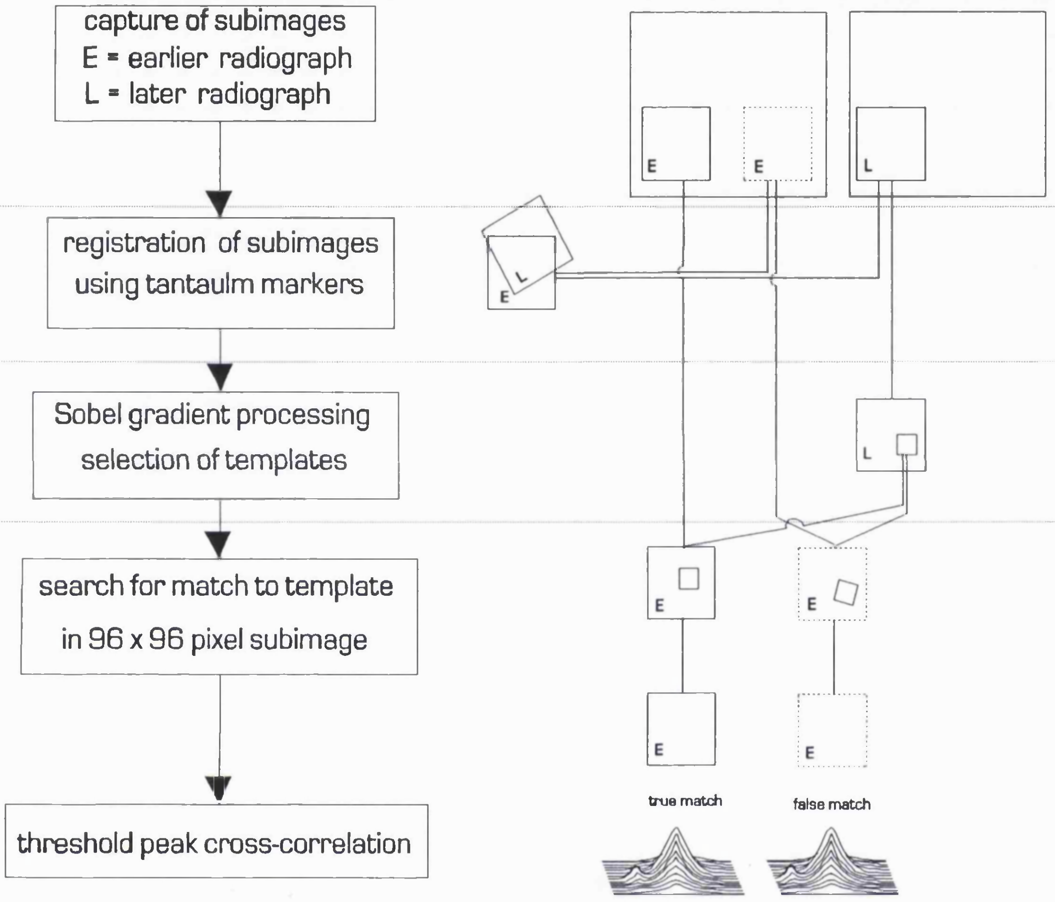


Figure 10.5 Flow chart and plan of experiment to evaluate the detection accuracy at the system operating point.

detection accuracy in one data set over the other.

10.3.3.4 *Estimation of local accuracy*

The local accuracy of the detection system was assessed at the normal system operating point (threshold for detection set at $\rho = 0.68$ and $\rho = 0.75$ for mandible and maxilla respectively) using the templates employed in the detection accuracy experiments (above). The system was operated conventionally to detect the presence of a match for each template. As before, the first stage of the system was used to detect a match in the entire corresponding subimage (96 x 96 pixels) from the earlier radiograph. For each template the coordinates of the highest cross-correlation (above the threshold) were recorded.

Following the initial detection search, the template and subimage from the earlier radiograph were decorrelated and a second search conducted to locate the position of match. This second search was carried out over a region of 30 x 30 pixels centred on the point of highest cross-correlation, irrespective of whether or not this site represented a true-match. Where the highest cross-correlation occurred close to the margin of the search area it was necessary to continue the search outside the confines of the subimage. This was achieved by reading the original earlier image from the central memory and relocating the search subimage.

The coordinates of the highest cross-correlation in this second search were transformed to the coordinate system of the later radiograph based on the positions of the implants. The transformed coordinates were then subtracted from those marking the centre of the template and the difference between the "true" (implant-based) and estimated locations recorded. These differences constituted the errors of local accuracy. This procedure was repeated for both jaws in all subjects. The design of the experimental method is shown diagrammatically in Figure 10.6.

Statistical analysis.- The distribution of the errors was examined for normality in the two coordinate axes of the image using Shapiro-Wilk's test. The means, variances and covariances of the errors were calculated for each coordinate axis using standard statistical formulae (Jaech, 1985). Analysis of the error was performed using Cochran's test for the homogeneity of variances to determine if significant differences existed between the magnitude of the error variances within or between the mandible and maxilla. Separate Cochran's tests were conducted for each data set. The relationship between the errors in the two coordinate axes was examined visually using scatter plots followed by more formal examination using Pearson's product moment correlation coefficient. The possibility of systematic error in the location of matching sites was

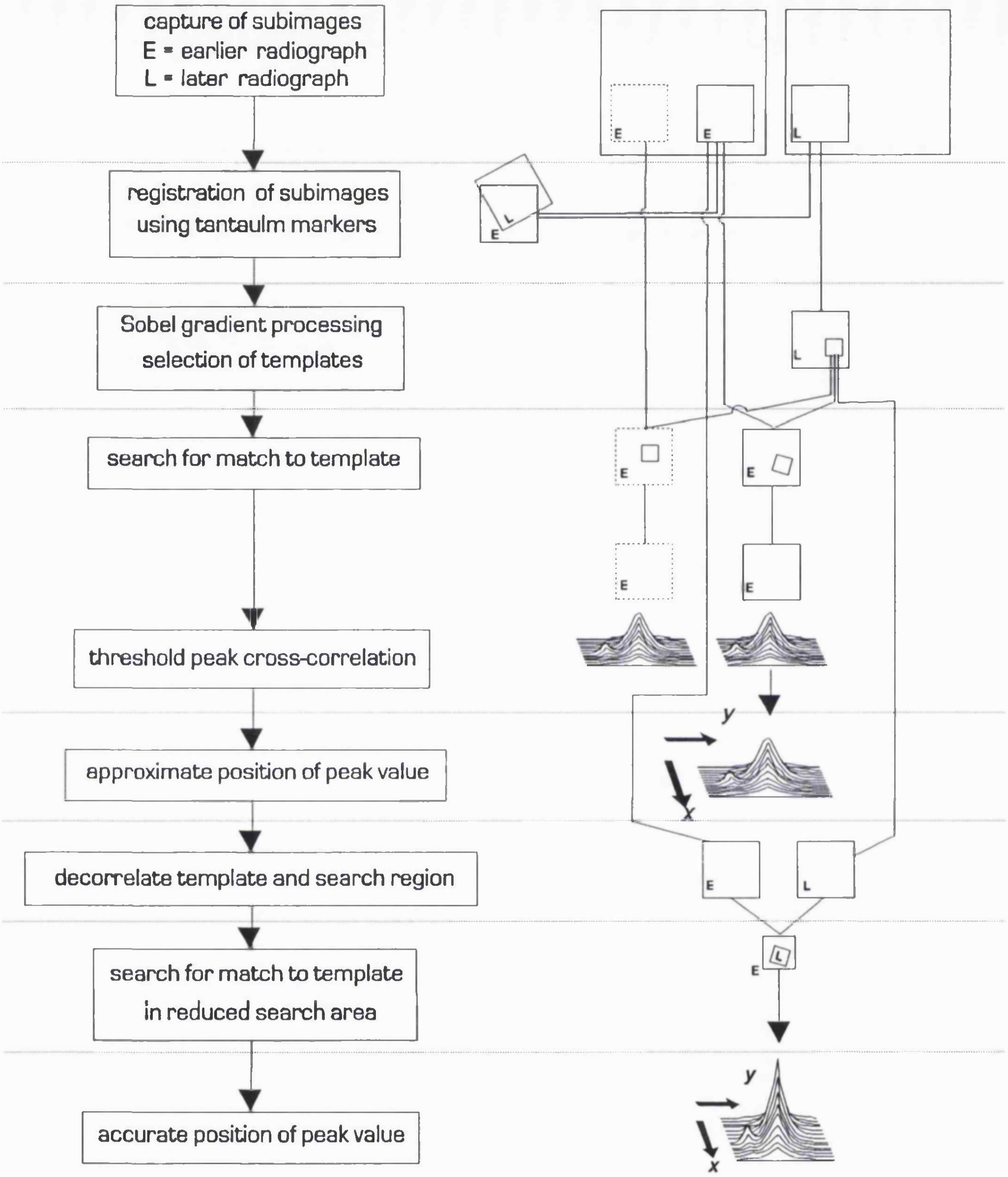


Figure 10.6 Flow chart and plan of experiment to evaluate local accuracy.

examined by testing the mean errors for the statistical significance of their differences from zero using one-sample *t* tests.

10.3.4 **Results**

The results are presented in a series of Figures, graphs and Tables. For each experiment two sets of results are given, one for each data set.

10.3.4.1 *Detection accuracy of the system*

The mean ROC curves constructed from the two data sets (1a and 1b+2) are shown in Figures 10.7 to 10.10. The areas under the individual ROC curves are presented in Table 10.2.

Table 10.2 Areas under the individual ROC curves for each of five trials.

Trial number	1	2	3	4	5
Data set 1a:					
mandible	0.76	0.77	0.72	0.74	0.75
maxilla	0.67	0.63	0.67	0.68	0.73
Data set 1b+2:					
mandible	0.73	0.67	0.73	0.74	0.72
maxilla	0.67	0.63	0.71	0.68	0.70

Note: area under ROC curve is given as a proportion of the maximum area of unity.

The areas beneath the ROC curves were statistically significantly different from that expected by chance alone both for the mandible ($p \leq 0.05$) and for the maxilla ($p \leq 0.05$) from data sets.

Based on the Kruskal-Wallis test, the difference in detection accuracy between the data sets was not significantly different for either the mandible or the maxilla at the 5% level ($p > 0.05$). However, the detection accuracy was found to differ significantly between the mandible and maxilla within each data set: data set 1a.- $p \leq 0.01$; data set [1b+2].- $p \leq 0.04$.

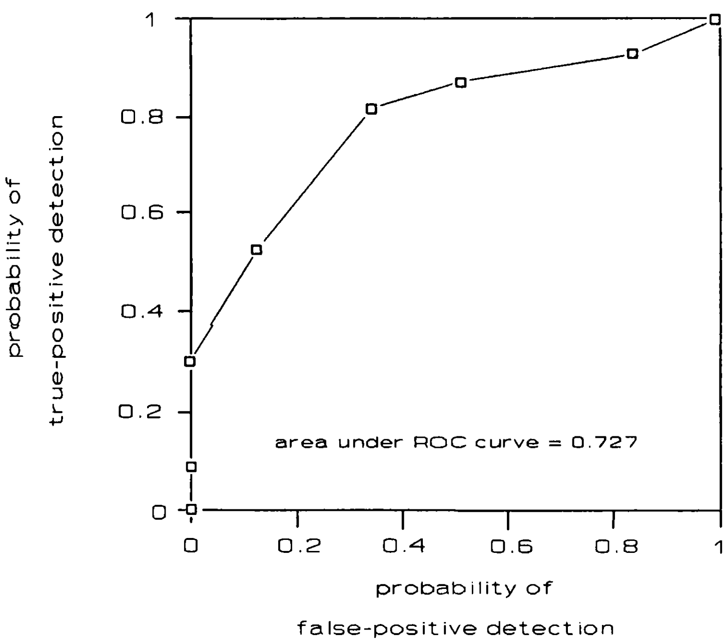


Figure 10.7 Mean ROC curve for reference sites in the mandible - data set 1a

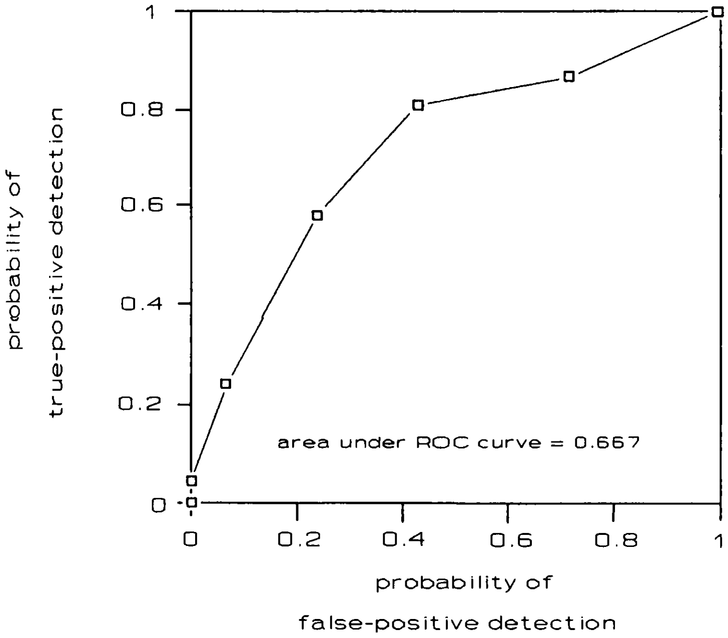


Figure 10.8 Mean ROC curve for reference sites in the maxilla - data set 1a

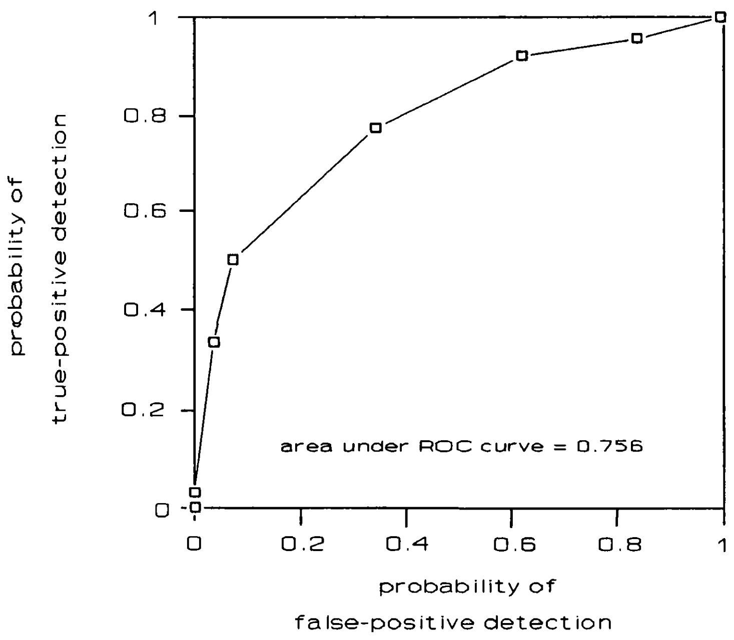


Figure 10.9 Mean ROC curve for reference sites in the mandible - data set 1b+2

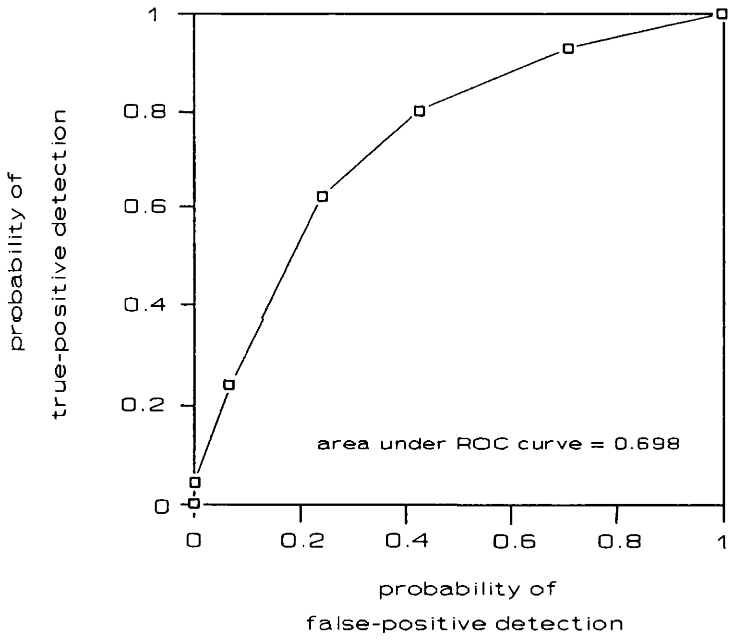


Figure 10.10 Mean ROC curve for reference sites in the maxilla - data set 1b+2

10.3.4.2 *Detection accuracy at the operating point*

The estimated true-positive and false-positive detection probabilities and the 95% confidence intervals for the true probabilities are shown in Figures 10.11 and 10.12. The χ^2 analysis of the true-positive/false-positive ratio recorded for the two data sets revealed no statistical significant for either mandible (χ^2=0.163, p=0.69) or maxilla (χ^2=0.135, p=0.71). However, unlike the system accuracy, the detection accuracy at the operating point was not found to differ between mandible and maxilla for either data set: data set 1a, χ^2=0.226, p=0.63; data set [1b+2], χ^2=0.214, p=0.64.

10.3.4.3 *Estimation of local accuracy*

The estimates of the means and variances of the local registration errors in the two coordinate axes are recorded in Table 10.3. The estimates of the error covariances and correlation coefficients between the two coordinate axes are given in Table 10.4 together with the 95% confidence intervals for the true correlation coefficients.

Table 10.3 Estimated means and variances of the local registration errors in the two coordinate axes.

	Mandible				Maxilla			
	$\hat{\mu}_x$	$\hat{\mu}_y$	$\hat{\sigma}_x^2$	$\hat{\sigma}_y^2$	$\hat{\mu}_x$	$\hat{\mu}_y$	$\hat{\sigma}_x^2$	$\hat{\sigma}_y^2$
Data set 1a	-0.06	0.05	0.52	0.30	-0.02	0.07	0.09	0.61
Data set 1b	-0.06	0.00	0.76	0.70	-0.02	0.08	0.14	0.54

Key:

$\hat{\mu}_x$ is the estimated mean error in the horizontal axis;
$\hat{\mu}_y$ is the estimated mean error in the vertical axis;
$\hat{\sigma}_x^2$ is the estimated error variance in the horizontal axis;
$\hat{\sigma}_y^2$ is the estimated error variance in the vertical axis.

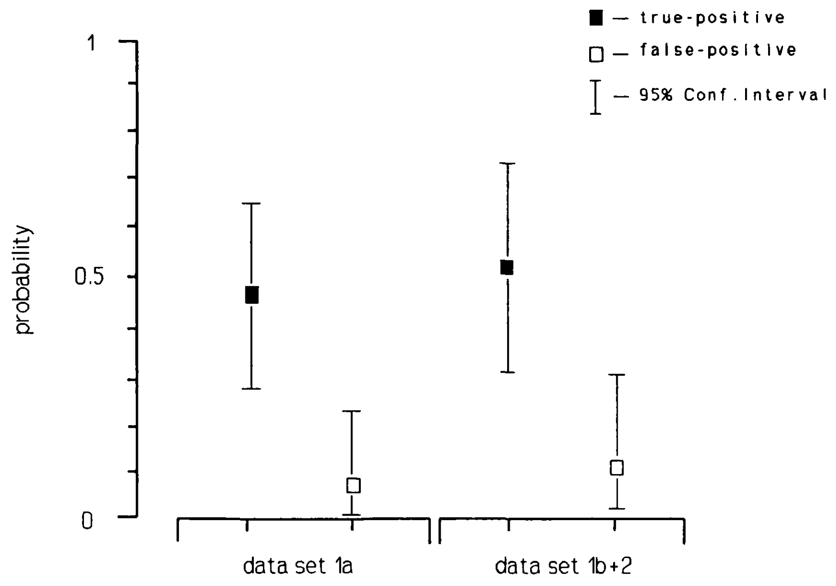


Figure 10.11 Estimated true-positive and false-positive detection probabilities and 95% confidence intervals at the system operating point for the mandible.

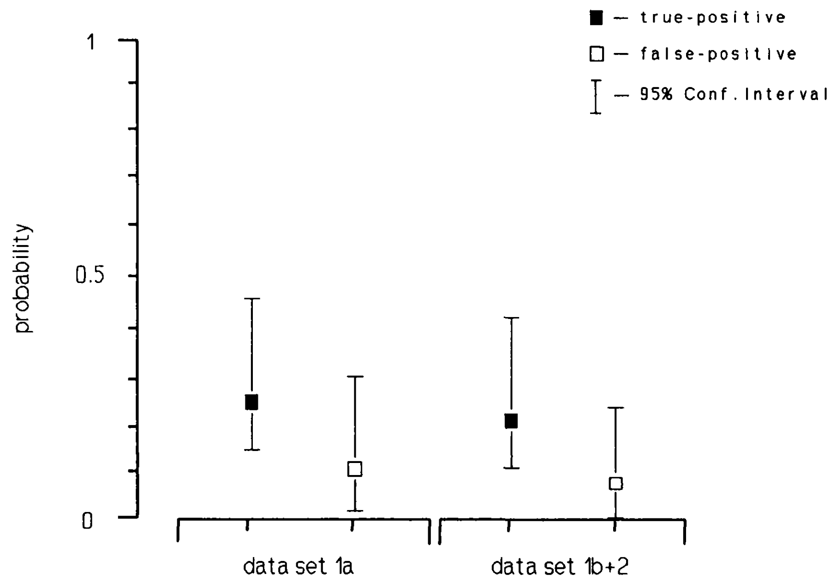


Figure 10.12 Estimated true-positive and false-positive detection probabilities and 95% confidence intervals at the system operating point for the maxilla.

Table 10.4 Estimated covariance and correlation between local registration errors in the two coordinate axes.

	Mandible			Maxilla		
	$\hat{\sigma}_{xy}$	$\hat{\rho}_{xy}$	95% CI ρ_{xy}	$\hat{\sigma}_{xy}$	$\hat{\rho}_{xy}$	95% CI ρ_{xy}
Data set 1a	0.40	0.05ns	-0.36 to 0.44	0.23	-0.61**	-0.81 to -0.27
Data set 1b	0.73	0.70***	0.44 to 0.85	0.27	-0.36ns	-0.65 to -0.02

Key:

$\hat{\sigma}_{xy}$ is the estimated covariance;

$\hat{\rho}_{xy}$ is the estimated product-moment correlation coefficient;

95% CI ρ_{xy} is the 95% confidence interval for the true correlation coefficient.

Levels of statistical significance:

ns	$p > 0.05$
**	$p \leq 0.01$
***	$p \leq 0.001$

Statistical analysis of the homogeneity of the error variances between the coordinate axes within and between the mandible and maxilla, revealed no significant difference for either data set: data set 1a.- $p>0.05$ (Cochran's statistic, D=0.400; critical value 0.404); data set [1b+2] $p>0.05$ (Cochran's statistic, D=0.357; critical value 0.392).

The means of the error distributions did not differ significantly from zero for mandible or maxilla in either data set, indicating a lack of registration bias (Figure 10.13).

Data set 1a, with 23 degrees of freedom.-

mandible: $\hat{\mu}_x$, p=0.30 (t=1.07); $\hat{\mu}_y$, p=0.20 (t=1.32).
maxilla: $\hat{\mu}_x$, p=0.47 (t=0.74); $\hat{\mu}_y$, p=0.15 (t=1.51).

Data set [1b+2], with 27 degrees of freedom.-

mandible: $\hat{\mu}_x$, p=0.30 (t=1.06); $\hat{\mu}_y$, p=0.93 (t=0.08).
maxilla: $\hat{\mu}_x$, p=0.52 (t=0.67); $\hat{\mu}_y$, p=0.11 (t=1.67).

Figure 10.13 Results of the one-sample *t* tests of the mean errors between the coordinate axes in mandible and maxilla in the two data sets.

10.3.4.4 *Summary of the results*

(1) The system detection accuracy was statistically significantly better than chance alone in mandible and maxilla for both data sets.

(2) The was no difference in the detection accuracy of the system between the data sets.

(3) The system detection accuracy was significantly higher in the mandible than in the maxilla for both data sets.

(4) The true-positive/false-positive detection ratios were similar for mandible and maxilla at the respective system operating points.

(5) There was no significant difference in the local accuracy for the two coordinate axes within or between the data sets.

(6) No significant registration bias was found in either coordinate axis for mandible or maxilla in either data set.

10.3.5 **Experimental investigation of detection sensitivity to rotational image distortion.**

In addition to the tests of accuracy a further series of tests were conducted to establish the sensitivity of the system to affine distortion in the images to be matched.

10.3.5.1 *Design of the experiments*

The assessment of the sensitivity of the system to affine distortion requires multiple radiographs of the same bone at different rotational alignments to the imaging system. Because it was not practical to obtain replicate radiographs of living subjects with tantalum markers an alternative method was employed. This involved the use of multiple radiographs of two human skulls with metallic markers secured to the surface of the mandible. The effect of growth changes could obviously not be examined by this method, however, the absence of such changes allows an upper limit to be placed on the detection precision for unchanging bony detail in the presence of rotational misalignment.

10.3.6 **Materials and methods**

The test images for this experiment were obtained from the replicated lateral cephalometric radiographs of the two dried human skulls in data set 5. Five radiographs were recorded of each skull at different rotations around the vertical axis perpendicular to the trans-porionic axis of the cephalostat (Figure 10.14).

To ensure the mandible remained in the defined position relative to the cephalostat each skull was radiographed in an inverted position. Films were recorded to provide an image quality generally similar to that obtained with radiographs of living subjects. The films were exposed at 50 mA and 65 kVp with an exposure time of 0.07 seconds using a 0.5 mm effective focal-spot size. Although it is possible to simulate the effects of contrast reduction from the investing soft tissue, pilot investigations using water-filled polythene pouches indicated major difficulties in controlling the orientation and stability of the skull within the cephalostat. Consequently, simulation of investing soft tissue was not undertaken.

Prior to recording the radiographs 2 mm diameter metal spheres were cemented to the mandibular body to act as registration markers: one in the mid-line between the

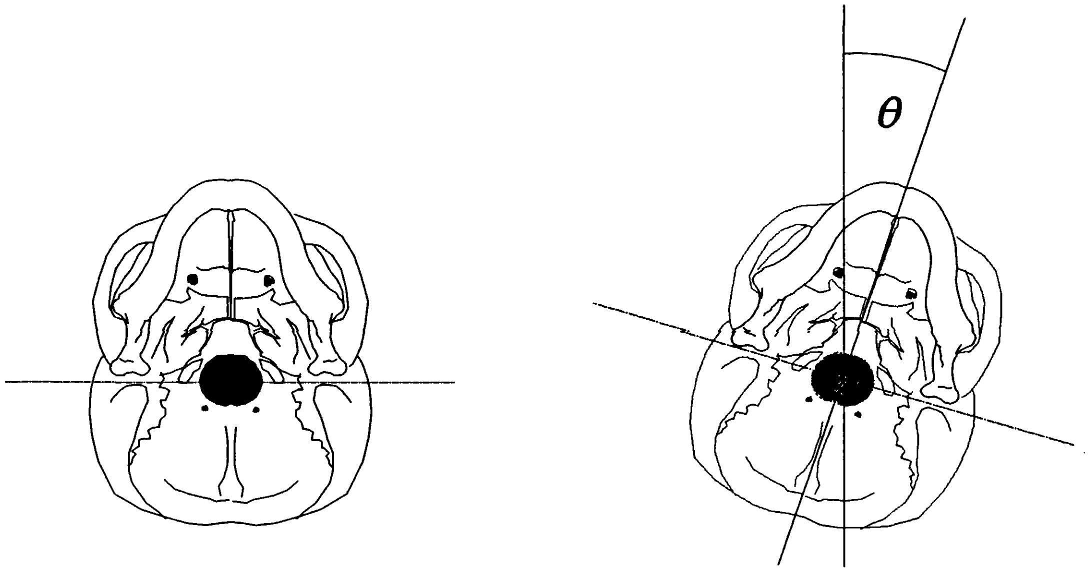


Figure 10.14 Diagram of the skull (basilar view) showing rotation (θ) about the vertical trans-porionic axis.

central incisors, one on each side over the first molar roots. To enable right and left sides to be distinguished the spheres on right and left sides were of different sizes: the right and left markers were 1 mm and 2 mm diameter respectively.

The initial exposure was made with the (inverted) skull in *norma lateralis* while subsequent exposures were made at 1° increments away from the film. This was achieved by rotating the cephalostat about the mid-porionic axis without removing the skull from the supports.

The radiographs were converted to digital format as described for the detection accuracy experiments (above) with the exception that subimages were only recorded for the mandible. A single subimage was captured from the premolar region of the 0° radiograph and processed to locate potential templates as described in chapter 9. Six templates were selected from each subimage. As before, the spatial position of each template was determined randomly from the potential sites in the central 40 x 40 pixels of the subimage.

The regions which corresponded to the origin of the template data were located in the 1°, 2°, 3° and 4° radiographs using the metal markers as reference points.

However, because the principal structures in the template were at an unknown depth, two regions were selected as corresponding to the origins of the template data - in right and left sides of the mandibular body (Figure 10.15).

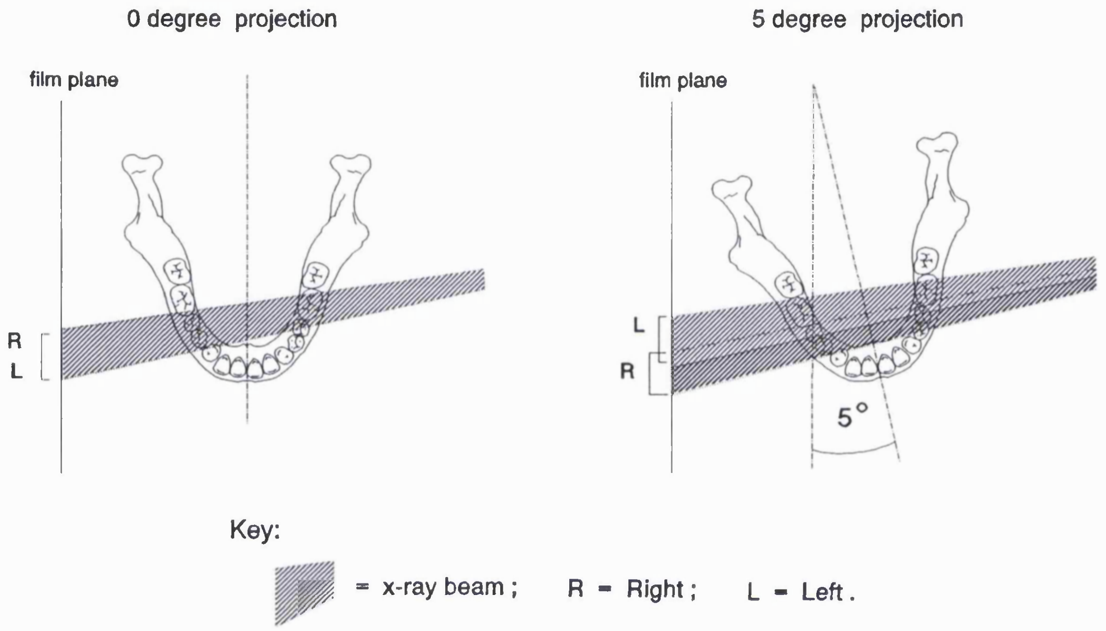


Figure 10.15 Diagram of the mandible showing the need to enlarge the search region as the skull rotates to encompass the possible origins of the template data in right and left sides of the mandibular body.

The detection system was used to search for a match for the template in each of these regions over an area of 25 x 25 pixels centred at the coordinates corresponding to the origin of the template. The detection search was conducted using only the first phase/stage of the system and detection was deemed to have occurred if the highest value of the cross-correlation exceeded the detection threshold (ρ=0.68) *in either or both search regions for each radiograph.*

Local detection accuracy was not investigated in these experiments because of the difficulty of valid interpretation where the depth of the structures in the template was unknown.

Statistical analysis.- The association between detection rate and the magnitude of rotational misalignment was analysed using Spearman's rank correlation. The null hypothesis, H_0: that the rank correlation coefficient was not different from zero was tested for statistical significance under the assumption that the templates depict independent data within and between the skulls.

10.3.7 **Results**

The relationship between the detection rate and rotational misalignment is shown in Figures 10.15 and 10.16.

Spearman rank correlation coefficient, r_s, indicated a strong inverse relationship between the peak correlation coefficient and the magnitude of rotation for both skulls: r_s = -0.94 (95% confidence interval, -0.99 to -0.33) for skull 1; and r_s = -0.83 (95% confidence interval, -0.99 to +0.21) for skull 2. A similar inverse relationship was also found between the proportion of sites correctly detected and the magnitude of rotation: r_s = -0.83 (95% confidence interval, -0.89 to +0.02) for skull 1; and, r_s = -0.93 (95% confidence interval, -0.99 to -0.23) for skull 2.

All the Spearman correlation coefficients found in this experiment were statistically significant at the 5% level ($p \leq 0.05$) except that for the proportion of sites and the magnitude rotation for skull 1. While this may appear surprising given the high correlation coefficient found (-0.83) the small number of readings has led to a wide 95% confidence interval which encompasses a correlation coefficient of zero.

The remaining results indicated strong inverse linear relationships which are significantly different from a correlation coefficient of zero, at the 5% level ($p \leq 0.05$).

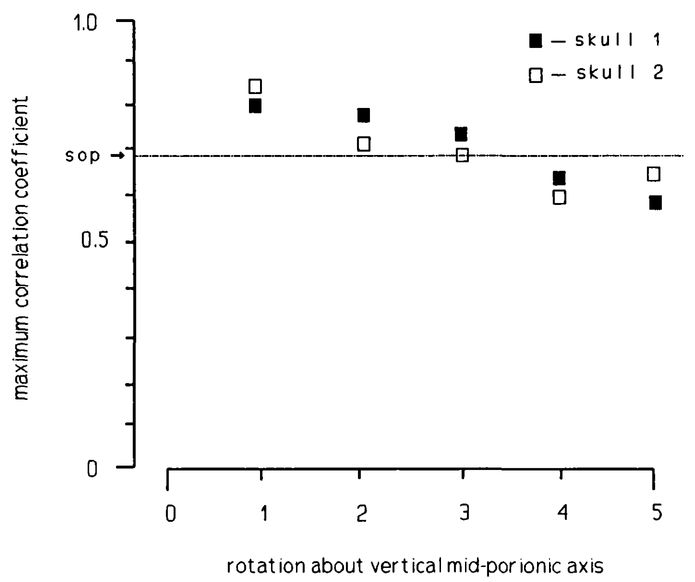


Figure 10.16 Graph of the peak correlation coefficient versus the rotation angle for mandibular reference sites in two dried skulls.

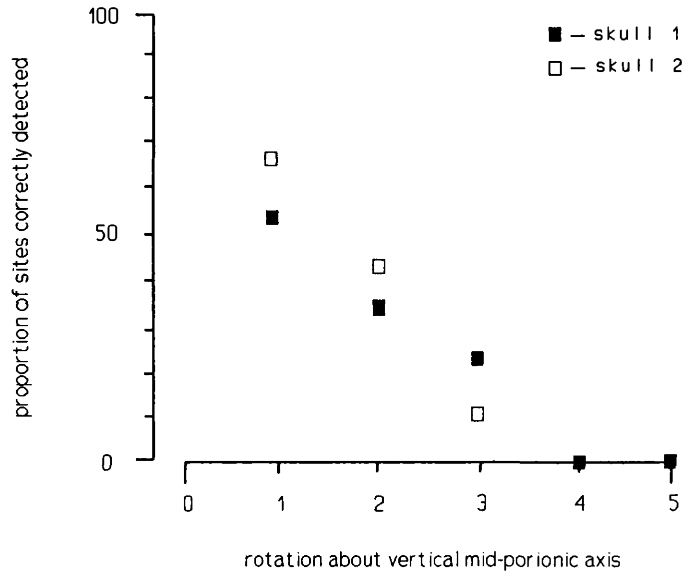


Figure 10.17 Graph of the proportion of sites correctly detected versus the rotation angle for mandibular reference sites in two dried skulls.

10.4 DISCUSSION

The series of experimental investigations described in this chapter were undertaken for two reasons: to see if the detection system worked effectively enough to be useful in practice; and to investigate the errors to which the detection system is prone, thereby placing quantitative bounds on the accuracy which can be expected in practice.

10.4.1 The methods used to test the detection system

In any experimental investigation designed to test performance, it is important to ensure that the tests are both appropriate and valid. In designing a test to achieve these aims for a detection system, there are two main considerations. The first, is the form of the test. The conventional approach is to examine the ability of the system to classify a set of data for which the true classification is already known. The second consideration is the size of the test sample which should be large enough to ensure plausible generalisations of the study results to the case population and large enough for valid applications of the statistical tests one proposes to apply (Swets and Picket, 1982).

These requirements presented particular difficulties because the necessary data could only be obtained from serial radiographs of growing bone where the true position of registration was known by reference to stable implanted markers. Only seven centres have undertaken controlled projection serial radiographic studies of normal children with implanted markers: the Royal Dental School, Copenhagen, Denmark; University of Bergen, Norway; Medical and Dental University of Tokyo, Japan; and four centres in the USA - University of Washington, University of California at San Francisco, University of Oregon, and Emory University. The number of children enroled in these studies varied from eight to just over two hundred, with a median of 24.

To obtain the samples for this study each of the centres was contacted (except Emory University dental school which had ceased to exist). Access was obtained to the radiographic records of a combined sample of 76 subjects and suitable radiographs were obtained from 54 subjects which form data sets 1 and 2.

Although the total sample is quite meagre it represents the largest single sample of normal children with stable implanted markers yet examined. Because of the ethical constraints which accompany both implant markers and exposure to ionising radiation, it is unlikely that larger samples will become available. Nonetheless, this sample is large enough, even when partitioned, to allow the application of standard statistical tests and is also large enough to allow credible generalisations of the results to the wider population, albeit with some caution. For example, based on a binomial distribution of the error, the system would need to make no false-positive detection errors in over two hundred and fifty trials to be reasonably certain that the "true" error rate was below two percent (Duda and Hart, 1973).

A further potential problem with the implant samples is the image quality of the duplicate films relative to originals. The faithful recording of distances is not generally a problem because the radiographs are copied by contact exposure. However, spatial resolution and contrast may be altered during duplication. Duplicate films may suffer up to a 20% reduction in contrast at the highest spatial frequencies examined in this study but with careful technique a reduction of only about 5% is readily attainable (Erales and Manson-Hing, 1979; Thunthy and Weinberg, 1982). For data set 2 the quality of the duplicates was directly verifiable from the optical density step-wedge at the margin of each film, but for data set 1 the quality of the duplicates could only be assessed subjectively.

In addition to the these potential problems there are four other possible criticism of the implant samples:

1) representativeness of the samples;
2) the limited age range (8 to 18 years);
3) excellent stability of projection geometry between exposures.

Of these, the stability of the projection geometry is potentially the most serious. The exceptional stability of the projection geometry arose from the need to verify the stability of the implants (that is, radiographs with differences in the positions of the implants, for whatever reason, were discarded). The consequence of this is that structures at different depths in the bone will have maintained the same positions in relation to the implants in each successive film, which may have provided an over optimistic assessment of the accuracy obtainable in practice.

The two competing aims of realism in the variation of projection geometry between films and the need to verify the stability of the implants cannot be accomplished by a two-dimensional image sequence recorded from a single perspective. Accordingly, the implant studies were supplemented by a simulation study employing dried skulls to examine the effects of varying projection between exposures.

10.4.2 Detection accuracy

The accuracy indicated by areas under the ROC curves demonstrates that the detection system is inherently capable of distinguishing correct and incorrect matches for small sections of bone image, at a rate better than pure chance. While this finding was not surprising it is an absolute prerequisite for any practical detection system. Of rather more interest was the finding that the system provided detection rates which were similar in both data sets. This finding was repeated throughout the experimental investigations and is of considerable importance because it provides clear evidence that the system is effective on data other than those used in its design. In addition, if the "true" performance of the system is bounded by the evaluations for the two data sets (1a and 1b+2), as suggested by Fukunaga (1990), then the system is clearly consistent in its detection ability.

10.4.3 The accuracy at the operating point

The accuracy at the operating point indicates how the system performs in a practical setting when the predetermined threshold is applied to the acceptance of a true match. The results would be expected to support the general findings for the system accuracy except for one important consideration. That is, when the system was used at the operating point only the highest cross-correlation value above the threshold was accepted; all other cross-correlation values were rejected. This point is of considerable importance because for each 15 x 15 pixel template a total of 8836 possible matches are assessed, and while there can be only one true point of match there are at least 8835 incorrect matches. Consequently, even if the probability of false-positive detection is very low, in such a large number of trails the detection of a false match above the threshold is almost inevitable, especially given the similarities in bone structure from region to region within the same bone. Selecting only the highest cross-correlation value reduces the possibility of false-positive detection.

While the estimated false-positive probabilities shown in Figures 10.11 and 10.12 present a true picture of the false matches detected by the system, it is important to remember that they represent false matches per search of 8836 possible matches. Accordingly, the probability of false-positive detection at any given point in the image was approximately $4 \div (8836 \times 8)$ or $p=0.000057$. The probability of true-positive detection was approximately $p=0.16$. Thus, while 84 in every hundred true matches were missed this was the price to be paid for suppressing the number of false-positive matches, since the detection of an apparent match where no match truly exists will generally be far more serious than a failure to detect a true match. This effect will be offset to some extent because of the large number of potential matches available. For example, a true-positive detection probability of 0.16 could, in theory, provide 1414 reference sites per 96 x 96 pixel subimage.

Examining Figures 10.11 and 10.12 appears to show a clear superiority for true-positive detection in the mandible over the maxilla. However, this was not supported by the χ^2 analysis. This is surprising, especially in view of the statistically significant difference found for the overall system accuracy. The explanation for this seems to lies in the limited information available from the single threshold rather than the entire ROC curve, coupled with the need to assume a parametric distribution in order to establish the confidence intervals. Nonetheless, this lack of difference between the accuracies at the operating points for the two jaws adds support to the decision to use different thresholds for mandible and maxilla.

10.4.4 **Estimation of local accuracy**

While the detection accuracy is of major importance in determining the presence and approximate location of stable reference sites, the effectiveness of the system depends on its ability to accurately locate positions of match. This depends on the mean and variance of the location errors.

The error variances were similar for both jaws and for both coordinate axes within each jaw. These results support the idea of circular symmetry of the errors and show that reference structures can be located with similar precision in both mandible or maxilla. However, the covariances and correlation coefficients of the errors within each jaw present a confusing picture of strong positive dependence between the coordinate axes in the mandible of data set 1b+2 and negative dependence between the coordinate axes of the maxilla in data set 1b. The remaining jaws showed no significant dependence at all. This baffling collection of associations probably arises from the combined effect of small sample size coupled with a correlation between the absolute value of the errors rather than between the signed (+'ve or -'ve) values. Recalculating the correlations independent of the sign of the errors provided correlation coefficients ranging from 0.30 to 0.78 but again only those for the mandible of data set 1b+2 and maxilla of data set 1a proved statistically significant. It is possible that larger samples would reveal a significant association between the errors in the remaining coordinate axes. The importance of this possible association is that large errors in one axis are likely to be matched by large errors in the other axis but in no particular direction.

It will be recalled that the system has been based around a maximum likelihood processor which is intended to locate matching structural detail with the minimum achievable error variance. This minimum is specified by a statistical expression known as the *Cramér-Rao inequality* (Whalen, 1971) and the lower limit of attainable error variance is referred to as the *Cramér-Rao Bound.* Because the detection system is based on a maximum likelihood processor, the errors of local accuracy should be close to the Cramér-Rao Bound. The Cramér-Rao Bound for the digital imagery used in this study was calculated using the formula provided by Ryan et al. (1980) and found to be approximately 0.015 mm^2. Although, the error variance found experimentally was quite low (range 0.09 to 0.70 mm^2) it was clearly much greater than the theoretical value. The most likely explanation for this is that the experimentally determined error variances included the registration errors from false-positive detection which is excluded from the calculation of the Cramér-Rao Bound. In addition, it is worth noting that although the Cramér-Rao inequality specifies a lower bound on the error variance it does not guarantee that this bound can be reached (Whalen, 1971).

Of broader significance to the practical application of the detection system, are the mean registration errors shown in Table 10.3. These mean errors were very small, the largest being less 0.1 mm, and none was significantly different from zero, indicating a lack of registration bias. Thus, even if the error variances were quite large, averaging the coordinates of several reference structures should provide a simple and effective method of reducing the mean error.

The number of reference structures whose coordinates must be averaged, will depend on the level of error which is tolerable. If the required accuracy, specified by the 95% confidence interval, is set at the dimensions of a single pixel (0.13 mm), then for the samples investigated here, an average of 29 structures would be required. Given that each 96 x 96 pixel subimage contains a potential 8836, 15 x 15 pixel templates of which 1414 (16%) are detectable, this level of accuracy appears readily attainable.

In the evaluation of a new method it is usual to analyse the findings in the context of the existing body of knowledge. This has not been possible here because of the lack of comparable investigations. However, the local accuracy found for the present detection system compares favourably with that for radiographic stereophotogrammetry (a three-dimensional system which employs tantalum implant markers rather than natural reference sites) where the error variance for linear measurements was approximately 0.04 mm^2 (Rune, 1980).

10.4.5 **Effects of rotational (affine) image distortion on detection accuracy**

Because conventional radiographs are two-dimensional projections of three-dimensional objects, rotations of the bone towards or away from the plane of the film between exposures leads to distortion between the two views. The importance of this type of distortion is that it makes detection more difficult since rotational changes in the alignment of structures relative to the x-ray source leads to changes in the spatial frequency content of the images. Consequently, because the detection system operates by assessing the spatial frequencies shared by the images to be matched, any comparison of spatial frequencies between images will be subject to error if their is a rotational distortion between the images.

The effect of this distortion is difficult to measure using radiographs of living subjects because of the need for multiple x-ray exposures at different rotational alignments. It was necessary therefore to employ a simulation using dried human skulls to assess these effects. The results of this simulation revealed a rapid decline in the peak cross-correlation as the rotation increased up to 4 degrees relative to the plane of the film. At rotations greater than 3 degrees no cross-correlation came above the detection threshold (Figure 10.16). The same decline was seen in the proportion of reference sites correctly detected with the proportion falling to zero at 4 degrees rotation (Figure 10.17). This very clear association between rotation and detection indicates the sensitivity of the system to minor changes in alignment of the bone relative to source and film.

The poor performance of the detection system at rotational misalignments greater than 4 degrees is not surprising as it is close to the 5 degrees rotational error at which standard cephalometric measurements fail (Grøn, 1960). Although this represents a clear limitation to the operation of the detection system, the level of rotational misalignment which can be expected between serial lateral cephalometric radiographs is usually well below 3 degrees. For example, Jeffcoat et al. (1987) found a mean rotational error of 0.32 (SD 0.1) degrees and a maximum of 1.50 degrees between replicate radiographs taken on the same day for six adults; and Julius (1972) found no rotational misalignments greater than 3 degrees for a group of 21 children followed over an average period of two years. Furthermore, Aouda and Johnson (1980) in an investigation of misalignment between serial cephalometric films stated that rotational errors "...are unlikely to normally exceed 4 degrees at which point the misalignment should be clinically apparent."

Thus, for the detection system to operate effectively the films must be recorded from very similar perspectives. While this requirement generally restricts the use of the system to large extra-oral radiographs, it is may be possible in future to overcome this problem for smaller intra-oral films or for films which record bony structures with only limited depth, by applying corrective distortions to the image data as proposed by Jeffcoat et al. (1984).

An importance issue which has not been specifically explored, is the relationship between the detection of reference sites and the time-lapse between recording the radiographs. This was not examined in the performance tests because of the limited time span between radiographs from the majority of subjects (only four subjects had radiographs separated by more than four years). While no clear indication of a difference in detection rates was seen in any of the experiments it is possible that a reduction in detection rate may occur as the time delay between radiographs increases. However, it seems more likely that specific periods of development are likely to be more important in reducing the available reference sites. For example, during the main phase of dental development and eruption. This aspect is explored in more detail in chapter 11.

The results of the performance tests suggest that:

1). Within the limitations imposed by two-dimensional radiographs, the detection system provides a means of detecting and locating persistent structural features in growing bone.

2). The positions of structural features located using this system, are consistent with the positions of stable implanted tantalum markers.

3). Although the system is sensitive to projection differences between radiographs, it is robust against the minor changes in projection likely to be encountered when the head is stabilised by a standard cephalostat or positioning device.

CHAPTER
11

THE APPLICATION OF THE DETECTION SYSTEM TO THE STUDY OF HUMAN INCISOR ERUPTION

11.1 INTRODUCTION

Knowledge of the variations of eruption and migration of the teeth is important in understanding the effects of orthodontic forces on the developing occlusion and in the planning of orthodontic treatment. This knowledge is particularly important in the case of the mandibular incisors whose labio-lingual position is used as the base line for the orthodontic re-positioning of the teeth.

However, despite their importance, surprisingly little is known about the pattern, direction or rates of movement especially during the intra-alveolar phase of tooth eruption (Berkovitz and Moxham, 1989). The reasons for this lack of understanding principally reside in the limitations of the methods presently available for the study of long-term changes in human hard tissues. Two basic methods of examination exist: histological; and radiographic. Histological studies of human tissues are necessarily cross-sectional, and while this type of study provides good qualitative material it allows only average quantitative changes to be determined. In addition, although general tooth movements can be deduced from observing average patterns of bone activity in histological sections, information about the distance and direction of tooth movement cannot be obtained in this way (Steedle and Proffit, 1985).

Radiographic analysis, on the other hand, provides the most satisfactory material for quantitative longitudinal comparisons. However, most radiographic studies have also been either cross-sectional or only analyzed on a cross-sectional basis, thereby limiting the information obtained to average timing or average patterns of movement (Schumaker and El Hadary, 1960; Darling and Levers, 1976). To acquire information on individual variations in timing, rate and direction of tooth movements requires long-term longitudinal study and very few investigations of this type have been undertaken (for example, Carlson, 1944; Burke, 1954; Björk and Skieller, 1972).

Unfortunately, longitudinal studies have been hindered by the lack of fixed reference points against which eruption can be measured. During the most rapid stages of movement, relatively accurate reference points may be provided by the occlusal surfaces of adjacent teeth or by bony landmarks. However, during the slower phases of movement the error associated with these landmarks is as great, or greater than the incremental changes being measured (Steedle and Proffit, 1985). Consequently, it is not possible to distinguish slow tooth movements without the use of reference markers which are fixed in the underlying bone (Reidel, 1972; Björk and Skieller, 1972).

Although orthodontically induced tooth movements have been observed relative to fixed markers in several studies (for example: Poulton, 1968; Björk, 1963; Mathews and Payne, 1982) only two such studies of natural tooth movement and eruption have ever been undertaken (Siersbæk-Nielsen, 1971; Björk and Skieller, 1972).

In the first of these studies, Siersbæk-Nielsen (1971) measured the eruption of mandibular and maxillary central incisors in a group of eight boys over a six year period centred around puberty. She also performed a more detailed examination of the variations in direction of eruption of the central incisors in one boy. Lateral cephalometric radiographs were used to record the movement of the teeth which were measured relative to tantalum implant markers. Consistent identification of the same incisor was achieved by restricting the examination to the most prominent incisor in each arch, which was located with the aid of study casts.

In the second study, Björk and Skieller (1972) also examined the migration and eruption of the teeth, relative to tantalum markers, around the age of puberty. Their sample consisted of twelve boys (eight of whom were from Siersbæk-Nielsen's sample) and nine girls. Although this study was longitudinal only average tooth movements were reported for the incisors and first molars.

Despite the advantages afforded by the use of tantalum implants it has not been possible to follow the pre-emergence or early post-emergence incisor movements from lateral radiographs because of the difficulty in distinguishing one tooth from another.

The method developed in the previous chapters provides a new mechanism for radiographic investigation of tooth eruption. In addition, the application of the detection system to the investigation of mandibular incisor eruption from serial lateral cephalometric radiographs should provide an exacting test of its capabilities because of the obscuring effects of multiple developing teeth and because of the limited bone available for selecting reference markers.

Consequently, the computer-based detection system was used to study mandibular incisor eruption from late intra-alveolar to pubertal eruption stages in a small group of untreated children.

11.1.1 **Aims and objectives**

There were two aims of this study: first, to allow a critical test of the detection system and secondly to investigate a problem which has eluded detailed analysis - the longitudinal eruption pattern of the mandibular incisors. These aims were elaborated into four objectives:

(a) To provide an exacting test of the detection system and thereby gain experience of its use in the analysis of multiple serial radiographs;

(b) To assess the ability of the detection system to locate and track the movement of individual teeth during active eruption in the presence of interfering signals from other superimposed teeth and bone;

(c) To evaluate the errors involved in this type of long-term investigation;

(d) To provide a pilot study into the course of human incisor eruption.

11.2 MATERIALS AND METHODS

11.2.1 Materials

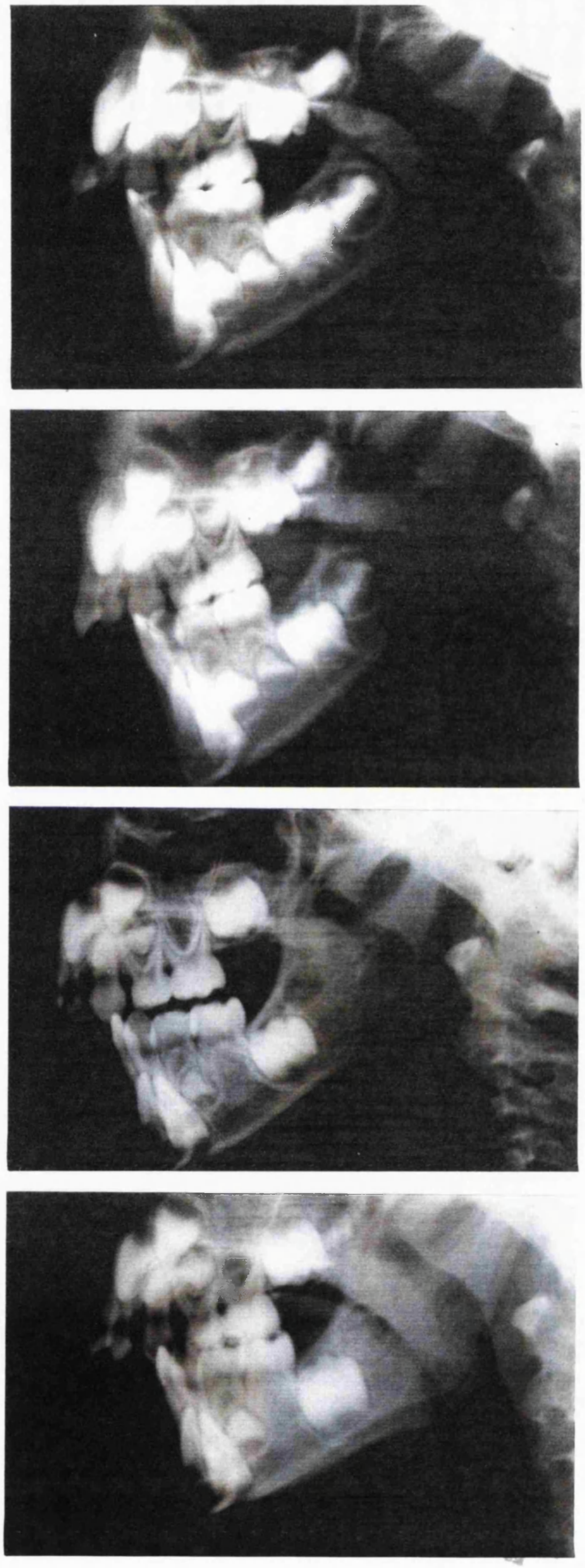
The experimental material comprised serial lateral cephalometric radiographs recorded at approximately six month intervals from age 4 to 7 years and annually thereafter to age 14 years for 11 children (7 girls and 4 boys). All radiographs were recorded with the same cephalostat and same source-film distance, providing a constant magnification of 1.07 at the mid-sagittal plane (MSP). All distances were corrected to absolute values at the MSP. The quality of the radiographic images was assessed as "very good" on the basis of subjective appraisal of the clarity of fine detail.

The subjects were drawn from a sample of 42 children whose dental and facial growth was monitored with serial radiographs during the period 1952 to 1970. Many of these children had brothers or sisters in the same sample. None of the subjects had received orthodontic treatment during the observation period. The subjects were selected only on the number and quality of their radiographic records and they could not be considered a truly random sample. The details of the cases and radiographic records are shown in Tables 11.1 and 11.2. Examples of the serial radiographs used in the study are shown in Figure 11.1.

Table 11.1 Details of the subjects whose radiographs were used in the study.

subject number	sex	Skeletal Pattern	ultimate incisor classification
1	f	II	II div.1
2	f	II	II div.1
3	f	III	III
4	m	II	II div.2
5	m	II	II div.2
6	m	II	II div.1
7	f	I	I
8	m	II	II div.1
9	f	I	II div.2
10	f	II	II div.1
11	f	II	II div.2

Key: f female; m male.



5½ years

5 years

4½ years

4 years

Figure 11.1 Examples of the serial lateral cephalometric radiographs used in this study.

Table 11.2 Distribution of radiographs used in the study.

subject Nº.	age (years) at which radiograph was recorded													
	4	4.5	5	5.5	6	6.5	7	8	9	10	11	12	13	14
1	●	●	-	●	●	●	●	●	-	-	●	●	●	●
2	●	●	●	-	●	●	●	●	●	●	-	●	●	●
3	●	●	●	●	●	●	-	-	●	●	-	●	●	●
4	-	●	●	●	●	-	●	●	●	●	●	●	●	●
5	-	●	●	-	-	●	●	●	●	●	●	●	●	●
6	●	-	●	-	●	●	●	●	●	●	●	●	●	●
7	-	-	●	●	●	●	●	●	●	●	●	●	-	●
8	●	●	●	●	●	●	●	●	●	●	-	●	●	●
9	●	-	●	-	●	-	●	●	●	-	●	●	●	●
10	●	●	●	-	●	-	●	●	●	●	●	●	●	●
11	●	-	●	-	●	●	●	●	-	-	●	●	●	●

Key: ● radiograph available; - no radiograph available.

11.2.2 Methods

The detection system as described in chapter 9 was used to locate:

a) stable natural reference sites in the bone;

and b) the crowns of individual incisors.

11.2.2.1 *Registration using natural reference markers*

The mandibular image from each radiograph was digitised using a pixel size of 0.13 mm using the experimental apparatus described in chapter 7. Prior to image capture, the maxillary plane of each radiograph was orientated parallel to the x-axis of the system to allow the x and y coordinates to correspond to approximate true horizontal and vertical axes respectively.

The detection system was used to locate reference sites from four regions away from the developing teeth (Figure 11.2). For each subject, one radiograph was selected to be the base-line reference image and this radiograph was used to establish the anatomical coordinate system for all other films. Pilot investigations had revealed that few (if any) bony reference sites would be available throughout the entire series of radiographs thereby making it impossible to directly register the oldest and youngest films. Therefore, a mid-range radiograph was used rather than using the earliest radiograph as the base-line reference. Thus, halving the time period over which reference sites needed to remain detectable.

Natural bony reference markers were located between the reference radiograph and each of the other radiographs as described in chapter 9, which were then used to register the images in pairs.

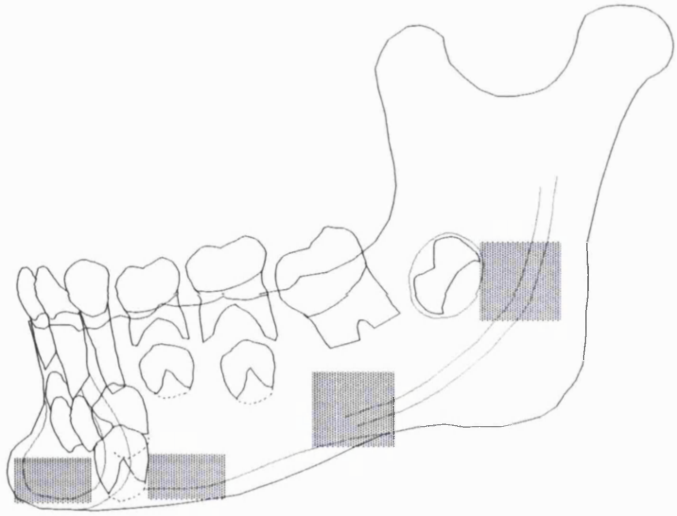


Figure 11.2 Sites of the anatomical reference markers. The anterior sites contained the principal reference markers; the posterior sites were used as an additional check on the rotational alignment of the images.

11.2.2.2 *Extraction of the reference signals for individual incisors*

In addition to detecting reference sites in the bone, the detection system was also used to locate the crowns of individual incisors. This was done in exactly the same way as detecting reference sites with the exception that each reference signal was the outline of the enamel crown of a particular incisor.

To provide accurate reference signals the crowns of individual incisors had to be selected and their images separated and cleaned of overlying, unrelated details. This was achieved by applying standard edge detection operators to the digitised images to locate

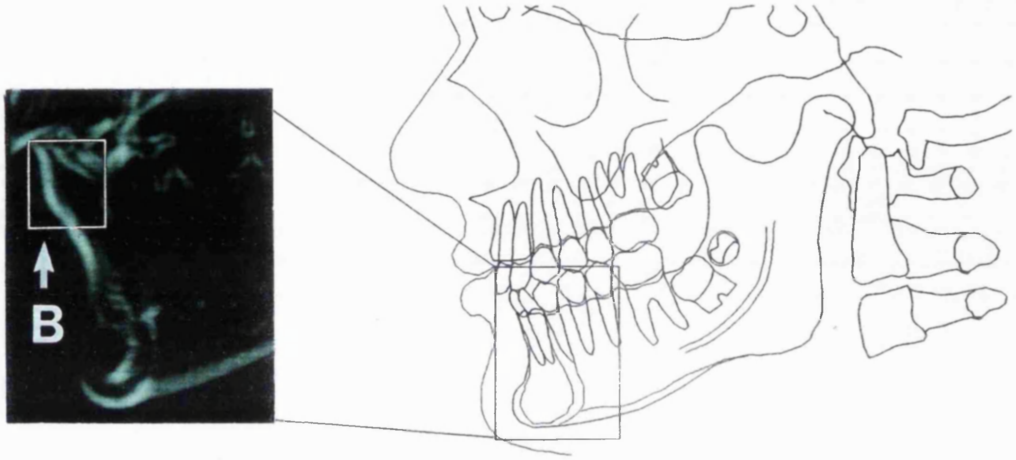


Figure 11.3 The location of the grey-level gradient (B) associated with the enamel outlines in the using a standard edge detection operator applied to the digitised image.

the grey-level gradient associated with the enamel outlines (Figure 11.3). Once the enamel outlines had been detected they were "cleaned" to remove high frequency noise by applying a *median filter* which replaced each pixel with the median value in the surrounding eight pixels (Gonzales and Wintz, 1987).

The enamel outlines were then located by eye using blue-green pseudo-colour to improve the visibility of the finest detail (Pratt, 1991) and the image of each external enamel outline was then extracted to provide each reference signal. Application of an interactively controlled threshold to the gradient image enabled the production of a binary valued mask of each tooth. The original image was then multiplied with the mask to remove all structure other than the outline of the enamel of the particular tooth. This final image was then used as the reference signal which was extracted and stored as a square template of 20 x 20 pixels.

11.2.2.3 *Tracking the eruption of the incisors*

The template containing the reference signal of each incisor was used to locate the position of the tooth in a second film by using the detection system to search for the cross-correlation peak.

Two searches were made, as in the conventional use of the detection system. The initial search was conducted using only the first two stages of the detection processor (that is, without employing the decorrelator). Having detected the presence and approximate location of each incisor crown in the second image the second search was conducted using the third and fourth stages of the detection system.

The impulse response of the decorrelator used in the location of the bony reference sites was unsuited to the high frequency details and strong spatial gradient associated with the outline of the enamel. Consequently, a new impulse response array consisting a first-order gradient operator, $\nabla_{x,y}$ was used instead (Pratt, 1991).

Once the crown of the incisor had been accurately located the x and y coordinates of the incisor tip were transformed to the coordinate system of the reference film and the process repeated for the next pair of images.

11.2.2.4 *Assessment of the eruption path.*- For each class of incisor (central and lateral) the transformed coordinates of the right and left incisor tips were averaged to provide a mean position from each successive radiograph. This was done to provide a single measurement relative to the MSP for each class of incisor which was unaffected by minor variations in the projection geometry from film to film.

The mean eruptive path for central and lateral incisors was then determined by a curve fitting procedure, thereby providing an estimate of the continuous path of the teeth from the series of sample points obtained from the radiographs.

To provide an accurate fit of the continuous curve (y) to the individual data points, a high-order polynomial was used, given by the following equation,

$$y = a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 \ldots a_{14}x^{14} \qquad \text{(Eq. 11.1)}$$

To provide a more accurate local fit to the oscillating post-emergence movements of the teeth, a second 14th order polynomial was fitted separately to this part of the eruption path.

The polynomial was fitted to the coordinate data points of each subject using the polynomial function of the ARCUS(12) data analysis program. The coefficients of the polynomial (a_0, a_1,....a_{14}) were determined by the computer program to provide a least squares approximation between the curve and the data points.

The amplitude of the post-emergence oscillation was estimated from the successive horizontal inflections of the eruption path determined from the fitted curve. The half period of oscillation was determined from the time delay between successive inflections of the eruption path. The curve fitting and estimation procedure are shown diagrammatically in Figure 11.4.

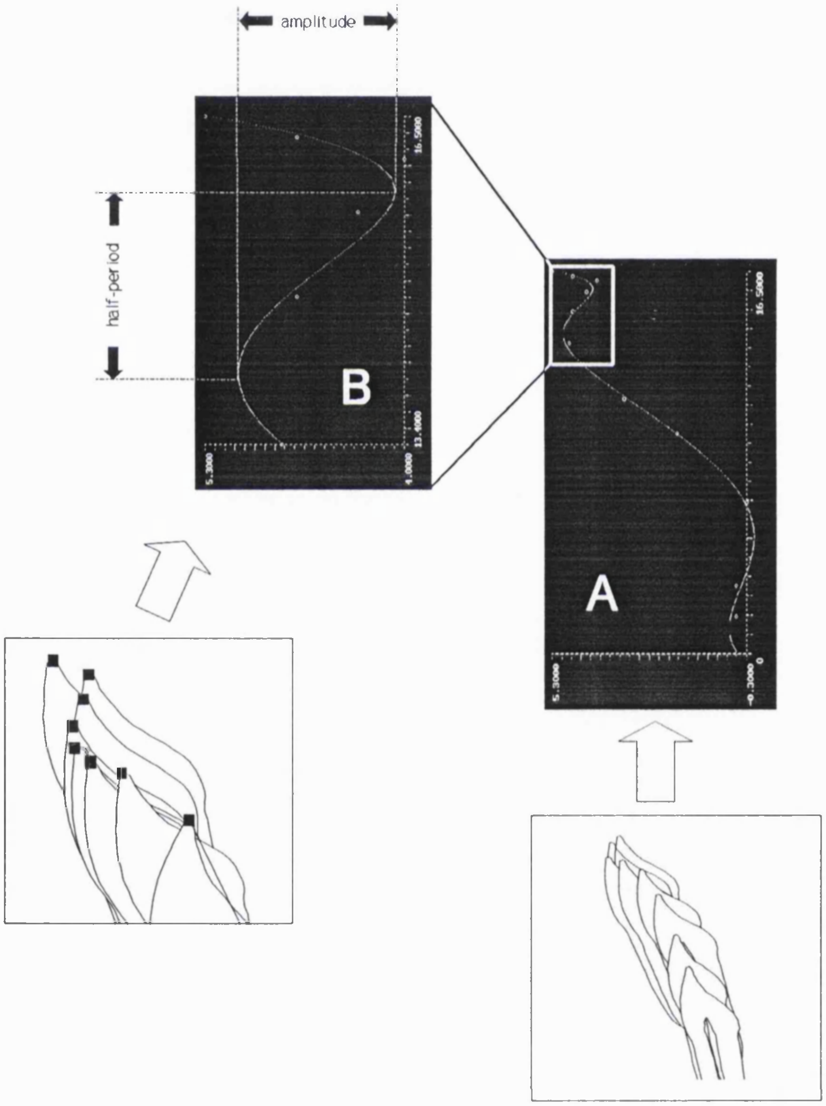


Figure 11.4 Diagrammatic representation of the curve fitting and estimation procedures.

11.2.3 Statistical methods

The eruption paths were determined individually for each subject and differences in the general form of the paths were not subjected to statistical analysis. The amplitudes and the periods of post-emergence oscillation were examined for normality using Shapiro-Wilk's test. Comparisons of the period and amplitude of oscillation between central and lateral incisors were conducted using paired *t* tests and possible associations between these parameters was examined using Pearson's product moment correlation.

11.3 ERRORS OF THE METHOD

In addition to the inherent errors of the detection system examined in chapter 10 an additional set of errors accompanies the application of the system to this type of longitudinal study.

The first of these errors arises because of the similarity in shape and orientation between right and left incisor antimeres. That is, the position and spatial frequency content of the different reference signals for each class of tooth are very similar. Thus, while the position of each tooth can be determined with great precision it is not easily possible to distinguish right from left.

This problem, coupled with the failure to maintain complete control over the projection geometry between the exposure of successive radiographs, leads to significant uncertainty in the relative positions of the same tooth in successive radiographs. This uncertainty arises independently of the accuracy of image registration.

To reduce the effect of these errors the coordinates of right and left lateral and central incisors were averaged to provide mean coordinates, referenced to the MSP, for each class of tooth.

The second of these errors arises because of the way in which the new position of the incisor tip at each stage was transformed back to the coordinate system established by the base-line radiograph. This type of approach not only leads to cumulative errors but also promotes the propagation of errors in the measurement of the total distance travelled or new direction of eruption at each stage. This form of error propagation is only rarely seen in conventional analyses of serial radiographs where successive films are generally superimposed directly on a single base-line film (Baumrind et al., 1987). Unfortunately there is no simple way to remove or reduce these errors.

11.3.2 Experimental determination of the method errors

To examine the errors of the experimental method two separate error studies were performed. First, the precision associated with detecting the change in incisor position coupled with registering the radiographic data via the reference film. Second, the precision of recording the oscillatory amplitude and period. Both studies employed paired data to provide estimates of the errors (Jaech, 1985).

The first study consisted of replicating the experimental procedure to determine the change in position of one randomly determined pair of incisors from the earliest and latest radiographs of each set. The change in mean incisor position was recorded for each subject. These data were then combined with the corresponding data from the main study to form the duplicate pairs. The differences between duplicate pairs were used to estimate the precision of the method and are shown in Table 11.3.

The second study consisted of replicating the experimental procedure for the oscillatory stage of post-emergence eruption for one randomly determined pair of incisors for each subject. The mean period and amplitude of oscillation were measured and recorded for each incisor pair. As in the previous error study, these data were combined with the corresponding data from the main study to form the duplicate pairs. The differences between duplicate pairs were used to estimate the precision associated with the measurement of the amplitude and period of oscillation. The results are shown in Table 11.3.

For both studies, the possibility of a systematic difference between double determinations was assessed using one-sample *t* tests, and the estimate of the precision and 95% confidence interval for the true precision were determined from the standard deviation of the differences in the two coordinate axes using the methods described by Jaech (1985).

11.3.2.1 *Results for the methods errors*

The results are presented in Tables 11.3 and 11.4.

Table 11.3 Means (mm) and variances (mm^2) of the differences for duplicate determinations of mean incisor position between earliest and latest radiographs.

x axis		y axis	
mean diff.	S_x^2	mean diff.	S_y^2
0.02	0.11	-0.03	0.07

Key: mean diff. is the mean difference between duplicate measurements;
s_x^2 is the variance of the differences in the horizontal axis;
s_y^2 is the variance of the differences in the vertical axis.

No significant systematic difference was found in either coordinate axis: x axis, p=0.28 (t=-0.61); y axis, p=0.51 (t=0.68).

The estimated precision, measured as the standard deviation of the error, was found to be:

x axis, (SD) = 0.33 mm (95% CI 0.16 mm to 0.41 mm).
y axis, (SD) = 0.19 mm (95% CI 0.13 mm to 0.33 mm).

Table 11.4 Means and variances of the differences for duplicate determinations of the period (years) and amplitude (mm) of oscillation.

period		amplitude	
mean diff.	S_p^2	mean diff.	S_a^2
-0.05	0.10	-0.07	0.12

Key: mean diff. is the mean difference between duplicate measurements;
s_p^2 is the variance of the differences between duplicate measurements of the period;
s_a^2 is the variance of the differences between duplicate measurements of the amplitude.

No significant systematic difference was found in the duplicate measurements for either: period, $p=0.64$ ($t=-0.48$); y axis, $p=0.50$ ($t=0.70$).

The estimated precision, measured as the standard deviation of the errors, were found to be:

Period of oscillation, (SD) = 0.22 years (95% CI 0.16 to 0.39 years);
Amplitude of oscillation, (SD) = 0.25 mm (95% CI 0.17 to 0.43 mm).

11.4 RESULTS

The findings are presented as a general description of the eruption pattern together with a more detailed metric analysis of the post-emergence eruption. In analyzing the findings no attempt was made to differentiate between the different classes of malocclusion or between the eruptive paths for males and females, because of the small sample size.

11.4.1 General pattern of eruption

The following description of the eruption paths of central and lateral incisors is based on the forms of the polynomial curves fitted to the mean (right and left) data points. A typical example of the mean eruptive paths and the relative positions of the incisor tips, are shown in Figure 11.5.

In all subjects, both central and lateral incisors showed a smooth path of intra-alveolar eruption which deviated lingually at, or just prior to alveolar emergence. This deviation was more marked for central than for lateral incisors (alveolar emergence was judged to have occurred when the incisal tip had crossed the bony alveolar outline). Both lateral and central incisors emerged lingual to the existing primary incisors.

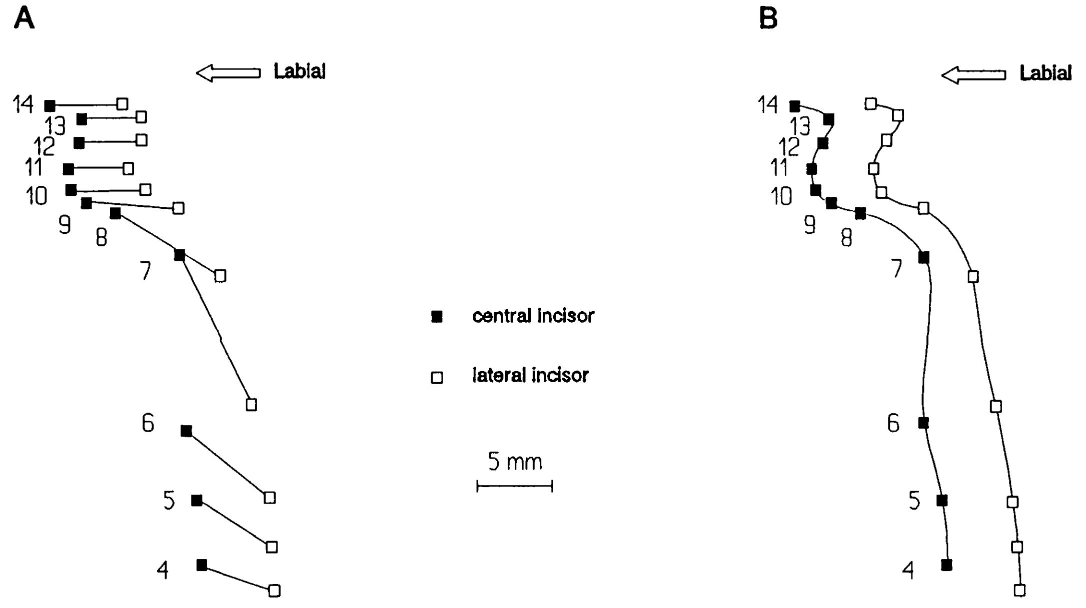


Figure 11.5 Example of: (A) the relative positions; and (B) the mean eruptive paths of the incisor tips in subject number 6. ■ marks the incisal tip.

Following emergence, both lateral and central incisors underwent a rotation of the crown, resulting in a more labially inclined path of eruption. This post emergence phase of eruption was rapid and was marked by a continued labial movement of the teeth as they approached the opposing arch. As the incisors moved vertically the labial movement reduced and then reversed. This process of labio-lingual movement persisted as an apparent oscillation of the crowns and continued even after contact was made with the opposing teeth[†] although the vertical movement slowed.

[†] The point of contact was judged radiographically.

Once the lateral incisors had erupted to the same vertical level as the central incisors the oscillation of central and lateral incisors synchronised. However, the amplitudes of the central incisors was equal to or greater than that of the lateral incisors in all cases. The oscillation was clearly evident for the central incisors in all subjects. For the lateral incisors an oscillation was evident in nine subjects but was equivocal in the two remaining subjects.

The maximum labial position of the incisor crowns was reached at about eleven years, the direction of horizontal movement then reversed. A further reversal back to a labial direction occurred in at about thirteen years. At the time of the final radiograph (14 years) the incisors in 9 of the 11 subjects were in a clear labial phase of movement.

During the post-emergence phase, the initial rapid vertical movement slowed as the moved labially up to about twelve years of age. Beyond twelve years the rate of vertical movement increased again in six of the subjects (4 girls, 2 boys). In the remaining subjects no change in vertical movement was observed up to 14 years.

11.4.2 **Analysis of the post-emergence eruption paths**

The results of the metric analysis of the post-emergence eruption in presented in Table 11.5.

The Shapiro-Wilk's tests revealed no significant deviation from a normal distributions at the 5% level for any parameter.

The periods of oscillation of central and lateral incisors were not significantly different p=0.89 (t=0.146) and showed a strong linear association within the subjects as indicated by the correlation coefficient, r=0.84 (95% CI 0.48 to 0.96) which differed significantly from zero (p=0.001).

The amplitudes of oscillation were generally less for lateral than central incisors and this difference was statistically significant, p=0.002 (t=3.6). Although, there was a significant linear association between the amplitudes of central and lateral incisors, r= 0.84 (95% CI 0.49 to 0.96; p=0.001).

There was no significant linear association between period and amplitude of oscillation for central incisors, r=0.35 (95% CI -0.31 to +0.79; p=0.66), however, there was a weakly significant association for lateral incisors between the period and amplitude, r=0.67 (95% CI 0.12 to 0.91; p=0.03).

Table 11.5 Period and amplitude of labio-lingual oscillation of the mandibular incisors.

	Period of Oscillation (years).		Amplitude of Oscillation (mm).	
subject number	central incisor	lateral incisor	central incisor	lateral incisor
1	4.6	4.4	2.4	1.9
2	4.7	4.7	2.2	1.8
3	4.0	4.4	2.4	2.4
4	3.9	3.8	2.2	0.8
5	4.5	4.1	1.4	0.5
6	4.9	4.9	2.7	2.5
7	4.6	4.6	2.5	2.6
8	4.4	4.6	2.3	1.2
9	5.1	5.2	3.2	2.7
10	4.6	4.5	1.4	0.7
11	4.4	4.4	2.3	2.2

Summary statistics:

mean period = 4.5 (standard deviation 0.3) years (for both central and lateral incisors).

mean amplitude (lateral incisors) = 1.8 (standard deviation 0.8) mm.

mean amplitude (central incisors) = 2.4 (standard deviation 0.5) mm.

11.5 DISCUSSION

This study was designed to serve two purposes: first, to allow a critical test of the detection system; and second, to investigate a problem which has eluded detailed analysis not only because of the difficulty of locating bony reference markers but because of the difficulty in unambiguously locating and following the crowns of individual teeth.

This type of investigation poses three major difficulties. First, the developing teeth obscure much of the bone limiting the area in which reference sites can be found. Secondly, the remodelling of the bone which accompanies eruption and the time-scale over which the teeth erupt means that all potential reference sites are likely to be lost during the study. Thirdly, distinguishing individual teeth and tracking their motion from the lateral view has hitherto been impossible because of the superimposition of adjacent teeth.

11.5.1 **Problems encountered in the study and the method errors**

11.5.1.1 *Reference site location*

The limited bone available from which to select the reference sites did not prove to be a major problem although the time-scale over which the teeth were followed made referencing all films to the earliest film impossible. That is, the same reference sites could not be used to register the entire series. Consequently, a mid-range film was used as the base-line reference. In theory, this should cause a reduction in the registration precision, however, the method errors were generally less than the local registration errors found in chapter 10. This apparent discrepancy probably arose because the estimates for precision were derived without replicating the radiographs and therefore exclude radiographic projection errors. Thus, although the precision appears very favourable it probably underestimates the true error and it is not directly comparable with the registration accuracy found during the evaluation of the detection system.

11.5.1.2 *Displacement of bilateral reference sites*

A particular problem with the radiographs used in this study was the variation in projection between successive films, which was not present to the same degree in the films used to evaluate the detection system. Consequently, it was not always possible to simultaneously register the images of bilateral reference sites which were situated at some distance from the MSP, that is, for reference sites in the mandibular body and ascending rami. Therefore, these sites were not used to provide the primary information to register the films but were used only as a check for gross rotational misalignment.

Because the principal eruptive movement was vertical, this restriction to reference sites close to the incisors should, in theory, have contributed a rotational component to the error which was greater horizontally than vertically. The relative magnitude of the horizontal and vertical errors supports this assumption (Table 11.3) but because of the wide 95% confidence intervals for the errors, the difference between horizontal and vertical errors was not statistically significant.

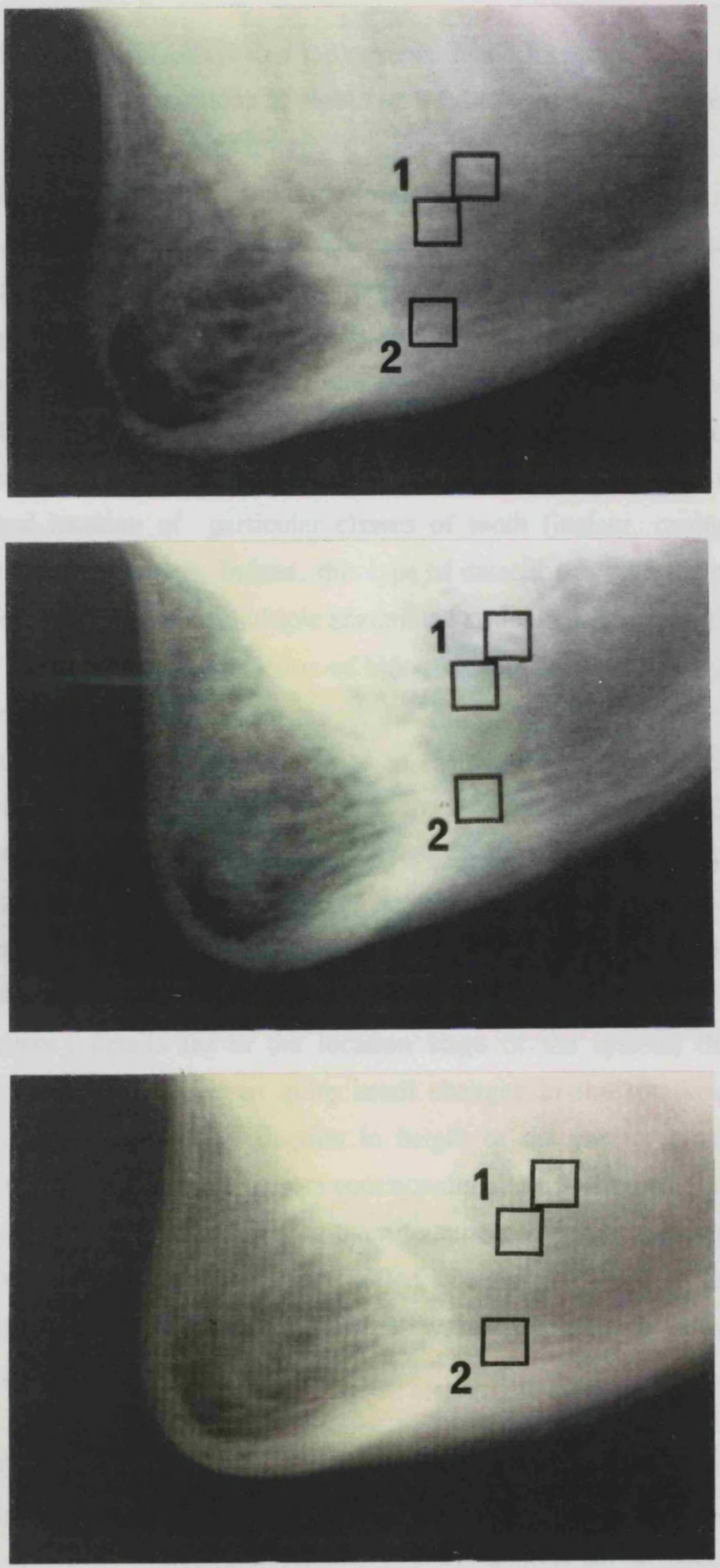
Natural reference structures.- An incidental finding which is likely to be of direct clinical interest was the recognition of two previously unreported natural reference structures in the anterior body of the mandible. These structures were located beneath the developing canine and premolars (Figure 11.6). The anatomical positions of the structures coincided with the neurovascular canal which perforates the mandibular cortex inferior to the developing canine and first premolar (Carter and Keen, 1971); and the fundic bone remnants left by the erupting canine and first premolar teeth. Throughout the study the region of the neurovascular canal provided reference details, which remained stable within the limitations of the method of measurement. The fundic bone remnants only appeared following the most rapid period of intra-alveolar eruption (8 to 9 years). These structures persisted in almost all subjects until the final radiograph at 14 years.

These reference sites may prove useful for the clinical analysis of cephalometric radiograph during the mixed dentition stage when the conventional reference structures are either obscured by the developing teeth or are not available (Björk and Skieller, 1983).

11.5.1.3 *Locating the incisor crown*

Perhaps the most major problem encountered in this study was the difficulty in distinguishing between right and left incisors. Although the detection system was not primarily designed to detect and locate teeth, the crowns of the teeth represent stable hard tissue which should be locatable in the same way as unchanging reference sites in bone. In spite of this, some problems were encountered which rendered the study less successful than had been hoped.

Spurious cross-correlation detection peaks.- During the detection of each incisor the cross-correlation surface generally showed two principal peaks rather than the single peak seen with the bony references. The second peak resulted from the spurious detection of the contralateral incisor crown, because of the very similar crown forms of the two teeth. Although consistent location of the same tooth appeared possible minor



14 years

12 years

11 years

Key:

1 Fundic bone remnants left by the erupting mandibular canine and premolar teeth.

2 Neurovascular canal which perforates the mandibular cortex inferior to the canine and premolar teeth.

Figure 11.6 Examples of mandibular reference structures found during this study.

Overlay to page 208

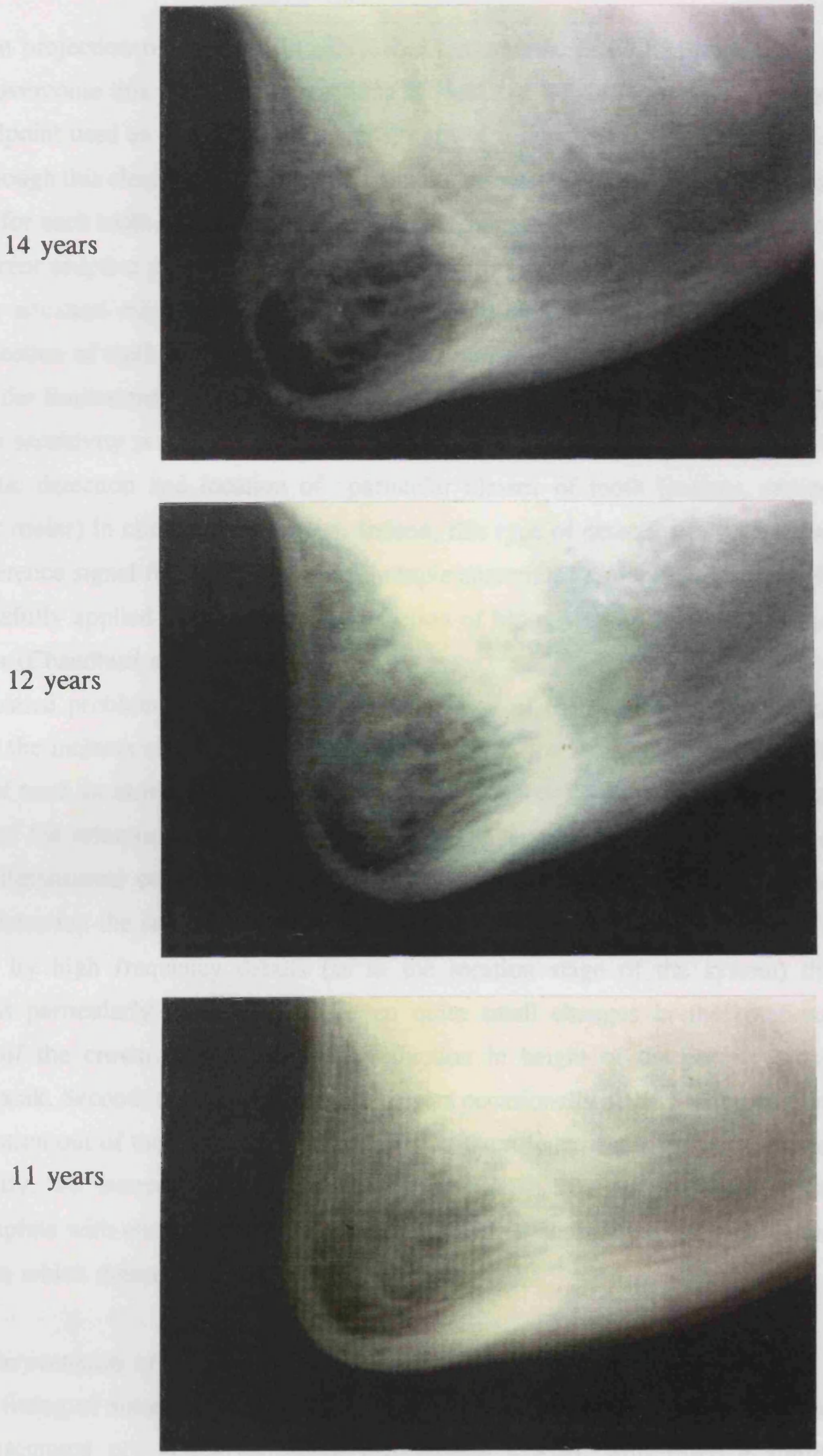


Figure 11.6 Examples of mandibular reference structures found during this study.

variations in projection of the teeth in successive radiographs made this impossible to verify. To overcome this problem the positions of right and left incisors were averaged and the midpoint used as the mean incisor position.

Although this clearly caused a reduction in the information about the exact pattern of eruption for each tooth, it provided the additional advantage of avoiding the distortion of the apparent eruptive path caused by differences in projection from film to film.

This situation might be improved if a threshold could be established for true-positive detection of teeth rather than of bone. Unfortunately, this is unlikely at present because of the limitations of spatial resolution in clinical radiographs. While this lack of detection sensitivity posed problems for this study it might be used to advantage in the automatic detection and location of particular classes of tooth (incisor, canine, premolar or molar) in clinical radiographs. Indeed, this type of detection system, using a single reference signal for the detection of multiple anatomical structures, has recently been successfully applied to the automatic detection of blood vessels in optical images of the retina (Chaudhuri et al., 1989).

A related problem was the change in orientation of the teeth which frequently occurred as the incisors erupted. This led to difficulties in the location (but not usually detection) of teeth in radiographs recorded some years before or after the film used as the source of the reference signal. The reason that this posed a problem was twofold. First, two-dimensional correlation-based detection systems are sensitive to rotational differences between the reference signal and object to be detected. Where the object is represented by high frequency details (as in the location stage of the system) this sensitivity is particularly marked. Thus, even quite small changes in the rotational orientation of the crown caused a marked reduction in height of the correct cross-correlation peak. Second, during eruption the incisors occasionally altered their positions by slow rotation out of the plane of the film, causing them to appear wider or narrower labio-lingually. To overcome these problems it was often necessary to replace the existing template with one derived from a radiograph which was chronologically closer to the one in which detection was being attempted.

11.5.1.4 *Interpretation of the fitted curve*

The fitting of smooth curves to sampled data points for the analysis of eruption has been attempted previously by Burke and Newell (1958), who fitted a simple exponential curve, and by Darling and Levers (1976) who used a more complicated exponential curve given by the so-called "Gompertz equation" (Darling and Levers, 1976). Both curves have well defined sigmoid forms which were clearly unsuited to the complicated curving form of the eruption path found in this study. However, the availability of modern computers provided an opportunity not available to previous investigators - the fitting of a high order polynomial curve which can be tailored more closely to the observed points.

Having established the data points the analysis of the mean eruption paths was conducted on the basis that the fitted curve was accurate. That is, that the fitted curve provided an accurate indication of the mean eruptive course of right and left incisors. While a 14th order polynomial curve provided a plausible fit to the data points it was inevitably accompanied by a residual error. It was apparent that the fit was poorer at points of rapid flexion: at the point of emergence; and at the oscillatory post-emergence stage. Hence, the need to use two high order polynomial curves to provide closer approximation to the data point.

Although, the fitted curves provided a close approximation to the data points defining the path of the central incisors, this was not always the case with the lateral incisors. Examination of the data points for the lateral incisors in nine subjects clearly revealed a cyclical pattern of positional change of the incisal tip. The remaining two subjects, this cyclical or oscillatory change, although barely discernable in the data points, has been provided by the fitted curve. This was clearly an effect of the curve fitting procedure.

Although the curves may have exaggerated some aspects of the eruptions paths in some subjects they reflect the most probable course taken during eruption as well as most probable in a mathematical "least-squares" sense.

Method errors for the period and amplitude of oscillation.- Despite the use of an automatic curve fitting procedure quite a large error of precision was found for the period of oscillation (95% CI 0.16 to 0.39 years) which appears to contradict the very similar amplitudes for central and lateral incisors within each subject. It seems likely that the explanation for this apparent contradiction is that although the curve was fitted automatically by computer, the period was estimated by eye from the same radiographs registered using the same reference sites. Consequently, the periods measured within each subject were highly correlated ($r=0.84$, $p=0.001$). A similar explanation is likely to apply to the periods of oscillation which were also highly correlated within each subject. For these reasons the apparent similarity of the amplitudes and of the periods for central and lateral incisors must be treated with some caution.

11.5.2 Discussion of the main study

11.5.2.1 *The general pattern of eruption*

Many attempts have been made to classify the stages of dental eruption using radiographic analysis. The most widely accepted seems to be Darling and Levers (1976) who divided eruption into 5 stages. However, Steedle and Proffit (1985) placed a further division in the second, eruptive stage to provided 3 pre- and 3 post-emergence stages.

Although only the late intra-alveolar to pubertal eruption stages were examined in this study, the general pattern of eruption was closer to that described by Carlson (1944) than to the more modern descriptions of Darling and Levers (1976) or Steedle and Proffit (1985). The difference lies principally in the *juvenile occlusal equilibrium stage* which was not seen in the present study. This stage is said to be a quiescent phase prior to the circumpubertal eruptive spurt. Although a diminution of vertical eruption was found, no arrest of occlusal movement was seen, as described by Darling and Levers (1974). In this respect the findings of this study support those of Carlson (1944), and Björk and Skieller (1972). This difference in findings is probably due to the cross-sectional nature of the method used by Darling and Levers (1976) which may have lacked the sensitivity to detect the slow occlusal movement revealed by longitudinal investigation. It is clear that when viewed from the lateral projection, the stages of eruption from the onset of root development fall into two clear groups: pre- and post emergence (Figure 11.7).

The intra-alveolar course of eruption was smooth, albeit with an increasing velocity as the tooth approached the oral cavity. The direction of eruption was along the long axis of the developing incisor. At, or just prior to emergence, however, there was a lingual deviation of the eruption path which was followed by the emergence of the incisor on the lingual aspect of its primary predecessor. It is important to note that this change in eruption path is not simply a consequence of providing a smooth mathematical curve, but is also clearly seen in the sampled coordinate data of the majority of subjects.

A possible explanation for this deviation in the eruption path is that the tooth was simply following the line of least resistance. The path is cleared ahead of the erupting incisor by the tissues of the dental follicle which may resorb the dense primary root more slowly than the thin edge of the alveolar bone. Support for this explanation may

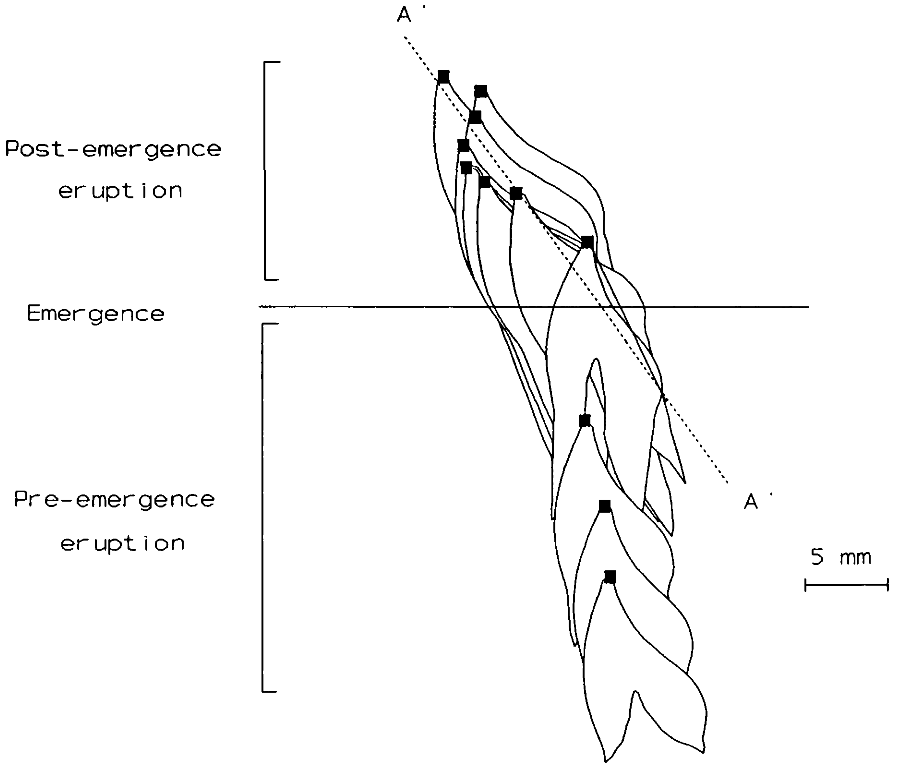


Figure 11.7 The basic pattern of eruption found in this study. A'-A represents the general line of labial motion of the incisors on which the labio-lingual oscillation was superimposed.

be found in the comparison of resorption rates for roots of primary teeth (Nanda, 1969) as compared with the resorption of alveolar bone (Cahill, 1969, 1970).

Following emergence, the incisor rapidly moved upward and labially with a marked change in inclination, possibly as a result of pressure from the tongue. However, a question of some interest is why the permanent incisors are more proclined and rapidly assume a more labial position than their primary predecessor, despite the absence of any obvious change in the balance of pressures between lips and tongue? In the light of the present study a possible answer is that as the tongue guides the crown labially the root, which is held in bone, is impeded which causes a rotation of the tooth about a centre located close to the alveolar margin. The ensuing change in inclination redirects the eruptive force labially. This added component of force together with the increased labio-lingual width of the labially inclined incisor, leads to a protrusion of the incisor tip beyond the position previously occupied by the more upright primary incisor (Figure 11.8).

It is possible that the lingual deviation of the crown at the point of emergence

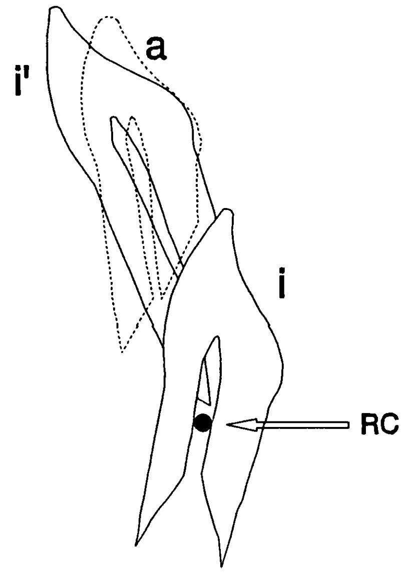


Figure 11.8 Diagram showing the relationship between the primary incisor (a) and the permanent incisor (i'). (i) indicates the position of the permanent incisor at emergence in relation to its centre of rotation (RC).

leads to greater proclination than would otherwise have been the case, by moving the centre of rotation lingually. Albeit that the lingual deviation was small, it may represent a previously unrecognised part of the mechanism which allows the permanent incisors to occupy the same basal bone as the primary incisors but without significant crowding.

Perhaps the most fascinating aspect of incisor eruption found in this study was the labio-lingual oscillation of the incisor crowns which was superimposed on a gradual labial drift of the incisors.

Following emergence, the eruption of the incisors was marked by a continued labial movement of the teeth as they approached the opposing arch. As the incisors moved vertically the labial movement reduced and then reversed. It is possible to attribute this initial reversal of labial movement to contact between the opposing incisors, by the action of the mandibular incisor tips being forced lingually against the cingulae of the upper incisors (Figure 11.9). However, this answer cannot explain the reversal in those cases with a large overjet, where contact with the opposing arch was not made during the observation period. In addition, the later return of the crowns to the original direction of motion cannot easilly be explained on the basis of the continued contact between opposing teeth alone.

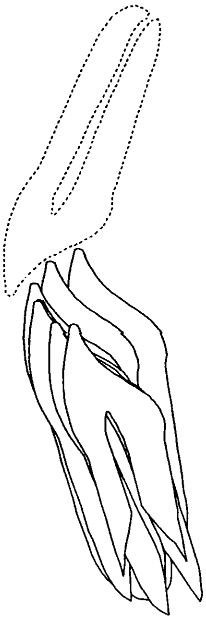


Figure 11.9 A possible explanation for the initial reversal of labial movement of the mandibular incisor as it is "guided" up the cingulum of the opposing tooth.

The arrest of the labial movement and the change in direction occurred at about 11 or 12 years of age. Although not compared in boys and girls, the narrow range of time at which the direction of eruption changed, indicates that the changes associated with the pubertal growth spurt are unlikely to be the cause. This tentative conclusion supports the previous observation by Saersbaek-Nielsen (1971).

The most obvious candidate for a controlling factor is the soft tissue balance between lips and tongue but it is difficult to see why the crowns should continue to oscillate unless the soft tissues balance is also oscillating. A possible answer lies in the mild oscillation of the mandibular corpus, which often occurs as part of a mandibular growth rotation (Björk and Skieller, 1983). A simple test of this hypothesis was conducted using cephalometric tracings of the mandible and anterior cranial base. The correlation coefficient between the direction of rotation of the mandibular corpus and the direction of movement of the incisors was 0.38 which is not significantly different from zero ($p>0.1$). Unfortunately, with such small for the correlation to be statistically significant it would need to be greater than 0.553 ($p \leq 0.05$). Consequently, this simple test of the hypothesis cannot be seen as conclusively rejecting a relationship between the mandibular rotation and incisor oscillation. It would be interesting to test this hypothesis with a larger sample of subjects.

Once the lateral incisors had erupted to the same vertical level as the central incisors their oscillation synchronised with that of the central incisors. It is interesting to compare the amplitudes of central and lateral incisors which were always equal to or greater in the central incisors. This finding is rather surprising as it seems to imply that the central incisors must move to the labial and then to the lingual of the lateral incisors, which is clearly not observed clinically. The explanation for this paradox may lie in the orientation of the teeth to the plane of view. The lateral incisors were set obliquely to the MSP and consequently only the antero-posterior component of the oscillation was visible. This may have led to underestimation of the amplitude and may also explain the difficulty in perceiving the oscillation in some subjects.

The finding that oscillation of the incisor crowns is widespread, if not universal, poses problems for one of the basic tenets of modern orthodontics - the stability of the lower labial segment. If further studies of incisor eruption corroborate these findings it may be necessary to re-evaluate the merits of the labio-lingual position of the lower incisors as the baseline for orthodontic re-positioning of the teeth.

CHAPTER
12

CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY

12.1 CONCLUSIONS

Although this study began with a clear problem of direct orthodontic and dental interest, the line of development dictated by the nature of the problem has meant the subject matter of this thesis covers areas far outside those normally encountered in dentistry. Many of these areas, such as signal detection theory and the manipulation of digital images, have become increasingly important in the analysis of clinical radiographs and radiographic systems during the past decade.

As a result of the development of the computer-aided detection system described in this thesis, it is possible for the first time, to detect valid reference markers for the registration and analysis of data from serial screen/film radiographs of the jaws, without the need for surgically implanted markers.

The practical detection system developed during this study should be considered as an initial crude implementation of a theoretical design which doubtless could be improved upon. Nevertheless, it is possible to conclude that:

1) The changes in bone structure seen in lateral cephalometric radiographs of growing bone, recorded several years apart, can be closely approximated by a simple additive noise model based upon a first-order Markov random field.

2). Matched filter based detection applied to serial radiographs of growing bone provides an accurate means for the detection of persistent structural features whose positions are consistent with stable implanted tantalum markers.

3). Many of the structural details which form the radiographically visible bony architecture of the mandible and maxilla remain unchanged and spatially stable over periods of several years.

4). Using the detection system developed in this study it has been possible to accurately locate persistent structural details in the mandible and maxilla in actively growing children with an error variance no greater than 0.7 mm^2 and without detectable mean bias. This latter finding indicates that increasing the precision by averaging the coordinates of several reference markers should also allow the accuracy of registration to be increased to levels previously unattainable even with implanted markers.

5). The effectiveness of the detection system requires the radiographs to be recorded with a rotational error no greater than 3 degrees. However, unlike subtraction systems commonly employed to register and analyse serial intra-oral radiographs, the system described here does not require the radiographs to be standardised to similar contrast and "brightness". However, pixel by pixel comparisons would require such standardisation.

In addition to these general conclusions, the application of the system to the investigation of mandibular incisor eruption revealed several previously unknown features of the pattern of incisor movement. While this study can only be regarded as a preliminary investigation because of the small sample size, nevertheless, some tentative conclusions do emerge:

6). The eruption paths of the mandibular incisors deviate lingually at or just prior to alveolar emergence. This is more evident in central than in lateral incisors.

7). The juvenile equilibrium phase of occlusion reported in previous cross-sectional studies probably does not exist as a separate stage. The arrest of vertical occlusal movement found in earlier studies presumably reflects the inability of cross-sectional investigations to disclose the slow occlusal movement revealed by the present longitudinal investigation.

8). The pattern of post-emergence incisor eruption is characterised by a slow labio-lingual oscillation of the crowns with an amplitude of approximately 1.8 mm and 2.4 mm for lateral and central incisors respectively. The period of oscillation is approximately 4.5 years for both central and lateral incisors.

9). The oscillations of central and lateral incisors synchronise following movement of the lateral incisors into the line of the arch.

10). As a result of the investigation of incisor eruption, two new natural reference sites for the clinical analysis of serial cephalometric radiographs have been provisionally identified. These are: the bony trabeculae surrounding the neurovascular canal which perforates the mandibular cortex immediately inferior to the developing canine and first premolar; and the fundic bone remnants left by the erupting canine and first premolar teeth.

The detection system is seen primarily as a research tool for analysing radiographs from existing growth studies, thereby maximising the information which can be gained from such studies. The applications of the system are in the area of facial hard tissue growth research including the eruption and development of the teeth. The monitoring of growth changes in children undergoing conventional orthodontic treatment is not considered a prime area of application. However, the monitoring of growth changes in children with severe cranio-facial deformities and the monitoring post-surgical changes and surgical relapse may prove feasible in time.

12.2 SUGGESTIONS FOR FURTHER STUDY

Further studies can be divided into three areas: Theoretical developments; practical improvements in the computer-based implementation of the detection system; and application of the system to problems of dental and skeletal hard tissue growth.

12.2.1 Theoretical development

The use of a more complex noise model explicitly incorporating distortion offers the most potential for improving the system and broadening its area of application. This is however, a complicated undertaking which will certainly require the adoption of several simplifying assumptions.

Recent theoretical and practical developments in the integration of matched filtering operations with higher-order statistics and polyspectra offers the possibility of detecting two-dimensional signals at very low signal-to-noise ratios (Giannakis and Tsatsanis, 1990). It would be very interesting to develop a detection system using these methods especially if they could be implemented using a neural network system.

12.2.2 Improvements in the computer-based implementation

Perhaps the most pressing need is to refine the computer-based implementation of the present detection system to allow it to run as a single program on an IBM compatible personal computer. In addition, it would be very helpful to develop a means to evaluate the "goodness" of matches produced by the system.

12.2.3 Application of the system to dental and skeletal hard tissue growth

The potential for application of the detection system in the field of hard tissue growth and pathological change is immense. Of particular interest from an orthodontic point of view is the interrelationship between skeletal growth rotation, posture and the development of malocclusion and such studies are planned.

Two further studies of dental eruption are planned and work to evaluate the system with intra-oral radiographs is in progress. This later application should prove of interest in the analysis of small changes in alveolar bone which could make possible the registration of unstandardised intra-oral radiographs along similar lines to the recently described technique of projective invariants (Dunn et al., 1993).

REFERENCES

Abramatic, J. F. and Faugeras, O. D. (1982) Sequential convolution techniques for image filtering, *IEEE Transactions on Acoustics, Speech and Signal Processing,* ASSP-30: 1-10.

Ackerman, R. J. (1976) Tooth migration during the transitional dentition, *Dental Clinics of North America,* 20: 661-670.

Agrawala, A. K. (1977) *Machine recognition of patterns,* IEEE Press, New York.

Andrews, H. C. (1970) *Computer Techniques in Image Processing,* Academic Press Inc., New York

Aouda, F. D. and Johnson, J. S. (1980) Possible effects of a misaligned lateral skull cephalometric radiograph on the values of routine angles used to aid clinical orthodontic diagnosis, *International Research Communications System; Medical Science,* 8: 175-176.

Arcese, A., Mengert, P. H. and Trombini, E. W. (1970) Image detection through bipolar correlation, *IEEE Transactions on Information Theory,* IT-16: 534-541.

Bajcsy, R. and Kovačič, S. (1989) Multiresolution elastic matching, *Computer Vision, Graphics and Image Processing,* 46: 1-21.

Barnea, D. I. and Silverman, H. F. (1972) A class of algorithms for fast digital registration, *IEEE Transactions on Computers,* C-21: 179-186.

Barrett, H. H. and Swindell, W. (1981a) *Radiological Imaging, Volume 1,* Academic Press, New York.

Barrett, H. H. and Swindell, W. (1981b) *Radiological Imaging, Volume 2,* Academic Press, New York.

Baumrind, S., Korn, S., Ben-Bassat, Y. and West, E. E. (1987) Quantitation of maxillary remodeling: 1. A description of osseous changes relative to superimposition on metallic implants, *American Journal of Orthodontics and Dentofacial Orthopedics,* 91: 29-41.

Baumrind, S., Ben-Bassat, Y., Korn, S., Bravo, L. A. and Curry, S. (1992) Mandibular remodeling measured on cephalograms: 2. A comparison of information from implant and anatomic best-fit superimpositions, *American Journal of Orthodontics and Dentofacial Orthopedics,* 102: 227-238.

Becker, A. and Karnei-R'em, R. M. (1992) The effects of infraocclusion: Part 1. Tilting of the adjacent teeth and local space loss, *American Journal of Orthodontics and Dentofacial Orthopedics,* 102: 256-264.

Berkovitz, B. K. B. and Moxham, B. J. (1989) Tissue changes during eruption, in: *Handbook of microscopic anatomy, Volume V/6: Teeth,* Oksche, A. and Vollrath, L. (eds)., Springer-Verlag, Berlin, pp. 21-71.

Bisgard, J. D. and Bisgard, M. E. (1935) Longitudinal growth of long bones, *Archives of surgery,* 1931: 568-578.

Björk, A. (1955) Facial growth in man, studied with the aid of metallic implants, *Acta Odontologica Scandinavica,* 13: 9-34.

Björk, A. (1963) Variations in the growth pattern of the human mandible: longitudinal radiographic study by the implant method, *Journal of Dental Research,* 42: 400-411.

Björk, A. (1968) The use of metallic implants in the study of facial growth in children: method and application, *American Journal of Physical Anthropology,* 29: 243-254.

Björk, A. (1969) Prediction of mandibular growth rotation, *American Journal of Orthodontics,* 55: 585-599.

Björk, A. and Skieller, V. (1972) Facial development and tooth eruption. An implant study at the age of puberty, *American Journal of Orthodontics,* 62: 339-383.

Björk, A. and Skieller, V. (1976) Postnatal growth and development of the maxillary complex, in: *Factors affecting growth of the midface,* McNamara, J. A. (ed)., Center for Human Growth and Development, Ann Arbor, Michigan, pp. 61-69.

Björk, A. and Skieller, V. (1977) Roentgencephalometric growth analysis of the maxilla, *Transactions of the European Orthodontic Society,* pp. 51-55.

Björk, A. and Skieller, V. (1983) Normal and abnormal growth of the mandible. A synthesis of longitudinal cephalometric studies over a period of 25 years, *European Journal of Orthodontics,* 5: 1-46.

Black, W. H. (1948) A method of roentgenographic cephalometric study of the growth changes in the facial bones of a dog, *M.S. Thesis,* Northwestern University.

Bogler, P. L. (1986) An analytical expression for correlation accuracy, *IEEE Transactions on Aerospace and Electronic Systems,* AES-22: 648-656.

Broadbent, B. H. (1931) A new x-ray technique and its application to orthodontics, *Angle Orthodontist,* 1: 45-66.

Brodie, A. G. (1949) Cephalometric roentgenology, history, technics and uses, *Journal of Oral Surgery,* 7: 185-198.

Buchanan, R. A. (1972) An improved x-ray intensifying screen, *IEEE Transactions on Nuclear Science,* NS-19: 81-85?

Buck, D. L. and Hodge, C. J. (1975) The within patient reliability of a three-dimensional cephalometric technique, *Archives of Oral Biology,* 20: 575-580.

Burke, P. H. (1954) Eruption analysis, *Transactions of the European Orthodontic Society,* 343-354.

Burke, P. H. and Newell, D. J. (1958) A photographic method of measuring eruption of certain human teeth, *American Journal of Orthodontics,* 44: 590-602.

Cahill D. R. (1969) Eruptive pathway formation in the presence of experimental tooth impaction in puppies, *Anatomical Record,* 164: 67-78.

Cahill D. R. (1970) The histology and rate of tooth eruption with and without temporary impaction, *Anatomical Record,* 166: 225-238.

Candy, J. V. (1988) *Signal processing: the modern approach,* McGraw-Hill, New York.

Carlson, H. (1944) Studies on the rate and amount of certain human teeth, *American Journal of Orthodontics,* 30: 575-588.

Carter, R. B. and Keen, E. N. (1971) The intramandibular course of the inferior alveolar nerve, *Journal of Anatomy,* 108: 433-440.

Cattermole, K. W. (1986) *Mathematical foundations for communication engineering. Volume 2: Statistical analysis and finite structures*, Pentech Press, London.

Chaudhuri, S., Chatterjee, S., Katz, N. and Goldbaum, M. (1989) Detection of blood vessels in retinal images using two-dimensional matched filters, *IEEE Transactions on Medical Imaging,* MI-8: 263-269.

Chen, Z. Q. and Paoloni, F. J. (1989) Electrical impedance images from measured data using an iterative reconstruction algorithm, *Electronics Letters,* 25: 38-40.

Claesson, G., Fredlund, P., Mühlow, A. and Selvik, G. (1978) Roentgen stereophotogrammetry for the evaluation of variations in liver volume and shape, *Acta Radiologica (Diagnosis),* 19: 423-432.

Cook, C. E. and Bernfeld, M. (1967) *Radar signals,* Academic Press, New York.

Cook, P. A. and Gravely, J. F. (1988) Tracing error with Björk's mandibular structures, *Angle Orthodontist,* 58: 169-178.

Cook, P. A. and Southall, P. J. (1989) The reliability of mandibular radiographic superimposition, *British Journal of Orthodontics,* 16: 25-30.

Cooley, J. W. and Tukey, J. W. (1965) An algorithm for the machine calculation of complex Fourier Series, *Mathematics of Computation,* 19: 297-301.

Crooks, L. E., Kaufman, L. Hoenninger, J. Arakawa, M., Watts, J. and Cannon, C. (1984) Spatial Resolution in NMR imaging, *IEEE Transactions on Medical Imaging,* MI-3: 51-53.

Cross, G. R. and Jain, A. K. (1983) Markov random field texture model, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-5: 149-163.

Dainty, J. C. and Shaw, R. (1974) *Image science: principles, analysis and evaluation of photographic type imaging processes,* Academic Press, New York.

Darling, A. I. and Levers, B. G. H. (1976) The pattern of eruption. In: *The eruption and occlusion of the teeth,* Poole, D. G. F. and Stack, M. V. (eds)., Butterworths, London, pp. 80-96.

Davenport, W. B. and Root, W. L. (1958) *An introduction to the theory of random signals and noise*, McGraw-Hill Book Company, Inc., New York.

De Coster, L. (1951) Hereditary potentiality versus ambient factors, *Transactions of the European Orthodontic Society,* 227-234.

Digby, P. G. N. and Kempton, R. A. (1989) *Multivariate analysis of ecological communities,* Chapman and Hall, London.

Doi, K. (1969) Weiner spectrum analysis of quantum statistical fluctuations and other noise sources in radiography. In: *television in diagnostic radiology,* Mosley, R. D. and Rust, J. H., (eds)., Aesculapius Publications, Birmingham, Alabama.

Dubreuil, G. (1913) La croissance des os des Mammifères. III. L'accroissement interstitial n'existe pas dans les os longs, *C. R. Societie Biologie, Paris,* 74: 935-937. Cited by Sissons, H. A. (1954), *Journal of Anatomy,* 87: 228-238.

Duda, R. O. and Hart, P. E. (1973) *Pattern recognition and scene analysis,* John Wiley and Sons, New York.

Duhamel, H. L. (1742) Sur le developpement et la crue des os des animaux, *Memoires Acadamie de Royale des Science,* 55: 354-370. Cited by Sissons, H. A. (1954), *Journal of Anatomy,* 87: 228-238.

Dunn, S. M. and van der Stelt, P. F. (1992) Recognising invariant geometric structure in dental radiographs, *Dentomaxillofacial Radiology,* 21: 142-147.

Dunn, S. M., van der Stelt, P. F., Ponce. A., Fenesy, K. and Shah, S. (1993) A comparison of two registration techniques for digital subtraction radiography, *Dentomaxillofacial Radiology,* 22: 77-80.

Egund, N., Olsson, T. H., Schmid, H. and Selvik, G. (1978) Movements in the sacroiliac joints demonstrated with roentgen stereophotogrammetry, *Acta Radiologica (Diagnosis),* 19: 833-846.

Engleke, K., Majumdar, S. and Genant, H-K. (1993) Impact of trabecular structure on marrow relaxation time, T_2, (abstract), *Calcified Tissue International,* 52: 174.

Erales, F. A. and Manson-Hing, L. R. (1979) A study of the quality of duplicated radiographs, *Oral Surgery,* 47: 98-104.

Faugeras, O. D. and Pratt, W. K. (1980) Decorrelation method of texture feature extraction, *IEEE Transactions on Pattern Analysis and Machine Intelligence,* PAMI-2: 323-332.

Feasby, W. H. (1981) A radiographic study of dental eruption, *American Journal of Orthodontics,* 80: 554-560.

Felsenberg, D., Nieber, V., Scheffler, S., Grampp, B. and Schneider, U. (1993) Bone determination by broadband ultrasound of the heel, (abstract), *Calcified Tissue International,* 52: 153-180.

Friede, H., Johanson, B., Ahlgren, J. and Thilander, B. (1977) Metallic implants as growth markers in infants with craniofacial anomalies, *Acta Odontologica Scandinavica,* 35: 265-273.

Frost, H. M. (1968) Tetracycline labelling in anatomy, *American Journal of Physical Anthropology*, 29: 183-196.

Frost, H. M. (1969) Tetracycline-based histological analysis of bone remodeling, *Calcified Tissue Research*, 3: 211-237.

Frost, H. M. (1973) *Bone remodeling and its relationship to metabolic bone diseases*, C. C. Thomas, Springfield, Illinois.

Fukunaga, K. (1990) *Introduction to Statistical Pattern Recognition.* 2nd Edition, Academic Press, London.

Gans, B. J. and Sarnat, B. G. (1951) Sutural facial growth of the Macaca rhesus monkey: a gross and serial roentgenographic study by means of metallic implants, *American Journal of Orthodontics,* 37: 827-841.

Garn, S. M., Silverman, F. N., Hertzog, K. P. and Rohmann, C. G. (1968) Lines and bands of increased density, *Medical Radiography and Photography,* 44: 58-89.

Geraets, W. G. M. and Van der Stelt, P. F. (1991) Analysis of the radiographic trabecular pattern, *Pattern Recognition Letters,* 12: 575-581.

Gerlot, P. and Bizalis, Y. (1988) Image registration: a review and a strategy for medical applications, in: *Proceedings of the 10th information processing in medical imaging conference,* De Graf, C. N. and Viergever, M. A. (eds)., Plenum, Utrecht, Netherlands.

Germano, G., Hoffman, E. J. and Dahlborn, M. (1990) An investigation of methods of pileup rejection for 2-D array detectors employed in high resolution PET, *IEEE Nuclear Science Symposium,* 2: 1263-1269.

Giannakis, G. B. and Tsatsanis, M. K. (1990) Signal detection and classification using matched filtering and higher order statistics, *IEEE Transactions on Acoustics, Speech and Signal Processing,* ASSP-38: 1284-1296.

Gill, G. G. and Abbott, L. C. (1942) Practical method of predicting the growth of the femur and tibia in the child, *Archives of surgery,* 45: 286-315.

Goldberg, M. H. and Sperling, A. (1981) Gross displacement of the mandibular canal: a radiographic sign of benign fibro-osseous bone disease, *Oral Surgery, Oral Medicine, Oral Pathology,* 51: 225-227.

Gonzalez, R. C. and Wintz, P. (1987) *Digital image processing,* 2nd edition, Addison-Wesley, Reading, Massachusetts.

Green, W. B., Bryant. N. A., Jepson, P. L., McLeod, R. G., Mosher, J. A., Selzer, W. D., Stromberg, W. D., Yagi, G. M. and Zobrist A. L. (1980) Analysis of multiple imagery at jet propulsion laboratory's image processing laboratory, *Optical Engineering,* 19: 168-179).

Greenwood, D. G. and Javid, S. F. (1989) DIMPL, Device Independent Image Processing Library, *Report No. 89/4,* Image Processing Group, University College London, London.

Grõn, P. (1960) A geometric evaluation of image size in dental radiography, *Journal of Dental Research,* 39: 289-301.

Gröndahl, H-G. (1992) Digital radiology in dental diagnosis: a critical review, *Dentomaxillofacial Radiology,* 21: 198-202.

Gröndahl, H-G. and Gröndahl, K, (1983) Subtraction radiography for the diagnosis of periodontal bone lesions. *Oral Surgery,* 55: 208-213.

Gröndahl, K. (1987) Computer-assisted subtraction radiography in periodontal diagnosis, *Swedish Dental Journal*, Supplement 50.

Haas, S. L. (1938) Growth in Length of the vertebrae, *Archives of Surgery,* 38: 245-249.

Harris, H. A. (1926) The growth of the long bones in childhood, with special reference to certain bony striations of the metaphysis and to the rôle of the vitamins, *Archives of Internal Medicine,* 38: 785-806.

Hausmann, E., Dunford, R., Wikesjo, U., Christersson, L., and McHenry, K. (1985) Progression of untreated periodontitis as assessed by subtraction radiography. *Journal of periodontal research*, 21: 716-721.

Helstrom, C. W. (1960) *Statistical Theory of Signal Detection*, Pergamon Press, New York.

Highleyman, W. H. (1962) The design and analysis of pattern recognition experiments, *Bell Systems Technical Journal,* 41, 723-744.

Hoyte, D. A. N. (1968) Alizarin Red in the study of apposition and resorption of bone, *American Journal of Physical Anthropology*, 29: 157-178.

Hulen, G. R. (1948) A method of roentgenographic cephalometry as combined with the use of metallic implants in studying growth of the mandible in the experimental animals: the dog, *M.S. Thesis,* Northwestern University.

Humphrey, G. (1864) On the growth of the jaws, *Philosophical Transactions of the Royal Society (Cambridge),* XI.

Hunt,B. R. (1976) Digital image processing, *Proceedings of the IEEE,* 63: 693-708.

Hunt, B. R. (1981) Nonstationary statistical image models (and their application to image data compression), in: *Image modeling*, Rosenfeld, R. (ed)., Academic Press, New York, pp. 225-238.

Hunter, J. (1770) Cited in: *The works of John Hunter,* Volume 4, J. F. Palmer, (ed)., Longmans Green, London (1837).

Isaacson, R. J., and Murphy, T. D. (1964) Some effects of rapid maxillary expansion in cleft lip and palate patients, *Angle Orthodontist,* 34: 143-154.

Ishizuka, R. (1981) The quantitative analysis of the mandibular trabecular x-ray image through two dimensional Fourier transform, *Nihon University Journal of Oral Science,* 7: 149-159.

Israel, H. (1968) Microdensitometric analysis for the study of skeletal growth and aging in the living subject, *American Journal of Physical Anthropology,* 29: 287-294.

Jaech, J. L. (1985) *Statistical analysis of measurement errors,* John Wiley and Sons, New York.

Jaffe, C. C. (1982) Lecture to American College of Radiology, *IEEE Transactions on Medical Imaging,* MI-1: 226-229.

Jain, A. K. (1981) Advances in mathematical models for image processing, *Proceedings of the IEEE*, 69: 502-528.

Jain, A. K. (1989) *Fundamentals of Digital Image Processing*, Prentice Hall, Englewood Cliffs, New Jersey, USA.

James, A. E., Anderson, J. H. and Huggins, C. B. (1985) *Digital image processing in radiology,* Williams and Wilkins, London.

Javid, S. F. and Greenwood, D. G. (1991) *YAM user manual*, Report No. 89/5, Image Processing Group, University College London.

Jeffcoat, M. K. (1992) Radiographic methods for the detection of progressive alveolar bone loss, *Journal of Periodontology*, 63:(supplement), 367-372.

Jeffcoat, M. K., Jeffcoat, R. and Williams (1984) A new method for the comparison of bone loss measurements on non-standardized radiographs, *Journal of Periodontal Research*, 19: 434-440.

Jeffcoat, M. K., Reddy, M. S., Webber, R. L., Williams, R. C. (1987) Extraoral control of geometry for digital subtraction radiography, *Journal of Periodontal Research*, 22: 396-402.

Jerome, C. P. and Gubler, H-P. (1991) Experimental determination of "The Law of Bone Remodeling" and effect of rat parathyroid hormone (1-34) infusion on derived parameters, *Calcified tissue International*, 49:398-402

Jordan, J., Bishop, P. and Kiani, B. (1989) *Correlation-based measurement systems,* Ellis Horwood, Chichester.

Julius, R. B. (1972) A serial cephalometric study of the metallic implant technique and methods of maxillary and mandibular superimposition, *M.S.D Thesis,* University of Washington.

Kailath, T. (1970) The innovations approach to detection and estimation theory, *Proceedings of the IEEE,* 58: 680-695.

Kashyap, R. L. (1981) Analysis and synthesis of image patterns by spatial interaction models, in: *Progress in Pattern Recognition*, Volume 1, Kanal, L. N. and Rosenfeld, A. (eds)., North-Holland Publishing Company, New York. pp. 149-186.

Kerningan, B. W. and Ritchie, D. M. (1978) *The C programming language,* Prentice-Hall, New Jersey.

Kittler, J. (1983) The template edge detector and its performance, in: *Issues in acoustic signal/image processing and recognition*, Chen, C. H. (ed)., NATO ASI series, Volume F1, Springer, Berlin-Heidelberg, pp. 296-397.

Knapp, C. H. and Carter, G. C. (1976) The generalised correlation method for the estimation of time delay, *IEEE Transactions on Acoustics, Speech and Signal Processing,* ASSP-24: 320-327.

Koskinen-Moffett, L., McMinn, R., Isotupa, K., and Moffett, B. (1981) Migration of the craniofacial periosteum in rabbits, *Proceedings of the Finnish Dental Society,* 77:83-88.

Krause, D. J., Steadman, J. W. and Williams, T. W. (1972) effect of record length on noise induced error in the cross correlation estimate, *IEEE Transactions on Systems, Man and Cybernetics,* SMC-2: 255-261.

Krogman, W. M. and Sassouni, V. (1957) *A syllabus in roentgenographic cephalometry,* Philadelphia Centre for Research, Philadelphia.

Lacroix, P. (1948) Le mode de croissance du perioste, *Archives de Biologie,* 59: 379-390.

Laitman, J. T. and Crelin, E. S. (1980) Tantalum markers as an aid in identifying the upper respiratory structures of experimental animals, *Laboratory Animal Science,* 30, 245-252.

Landau, H. J. (1967) Sampling, data transmission and the Nyquist rate, *Proceedings of the IEEE*, 55: 1701-1706.

Lee, M. M. (1968) Natural markers in bone growth, *American Journal of Physical Anthropology,* 29: 295-310.

Ling, A., Krile, T. Mitra, S. and Shihab, Z. (1986) Early detection of glaucoma using digital image processing. *Investigative Ophthalmology and Visual Science,* 27: (ARVO supplement) p. 160.

Litton, J. E., Holte, E. and Eriksson, L. (1990) Evaluation of the Karolinska new positron camera system: scanditronix PC2048-15B, *IEEE Transactions on Nuclear Science,* NS-37: 743-748.

Lorre, J. L. and Gillespie, A. R. (1980) Artifacts in digital images. *Proceedings of the Society of Photo-Optical Instrumentation Engineers,* 264: 188-194

Lynn, P. A. (1982) *An introduction to the analysis and processing of signals,* 2nd edition, MacMillan Press, London.

McGarty, T. P. (1974) *Stochastic systems and state estimation,* John Wiley and Sons, New York.

McGillem, C. D. and Svedlow, M. (1976) Image registration error variance as a measure of overlay quality, *IEEE Transactions on Geoscience Electronics,* GE-14: 45-49.

McGillem, C. D. and Svedlow, M. (1977) Optimal filer for minimization of image registration error variance, *IEEE Transactions on Geoscience Electronics,* GE-15: 257-259.

McWilliam, J. S. (1983) The effect of growth on the precision of subtraction superimposition, *Dentomaxillofacial Radiology,* 12: 61-69.

Mansfield, P. and Maudsley, A. A. (1977) Medical imaging by NMR, *British Journal of Radiology,* 50: 188-194.

Mathews, J. R. and Payne, G. S. (1980) Quantitative computerized analysis of lower incisor changes: a longitudinal implant study in man, *Angle Orthodontist,* 50: 218-229.

Mathews, J. R. and Ware, W. H. (1978) Longitudinal mandibular growth in children with tantalum implants, *American Journal of Orthodontics,* 74: 633-655.

Melsen, B. (1974) The cranial base. The postnatal development of the cranial base studied histologically on human autopsy material, *Acta Odontologica Scandinavica,* 32: Supplement 62.

Meyer, W. J. (1984) *Concepts of mathematical modeling,* McGraw-Hill, New York.

Middleton, D. (1960) *Introduction to Statistical Communication Theory,* McGraw-Hill, New York.

Miszalok, V., Seiler, T. and Wollensak, J. (1984) Quantitative evaluation of long-term fundus changes. *Investigative Ophthalmology and Visual Science,* 25: [ARVO supplement] p. 255.

Mitchell, D. L., Jordan, J. F. and Ricketts, R. M. (1975) Arcial growth with metallic implants in mandibular growth prediction, *American Journal of Orthodontics,* 68: 655-659.

Mostavi, H. and Smith, F. W. (1978a) Image correlation with geometric distortion. Part 1: acquisition performance, *IEEE Transactions on Aerospace and Electronic Systems,* AES-14, 487-493.

Mostavi, H. and Smith, F. W. (1978b) Image correlation with geometric distortion. Part 2: effect on local accuracy, *IEEE Transactions on Aerospace and Electronic Systems,* AES-14, 494-500.

Nanda, R. S. (1969) Root resorption of deciduous teeth in Indian children, *Archives of Oral Biology,* 14: 1021-1030.

Newell, J. C., Gisser, D. G. and Issacson, D. (1988) An electric current tomography, *IEEE Transactions on Biomedical Engineering,* BME-35: 828-833.

North, D. O. (1963) An analysis of the factors which determine signal-to-noise discrimination in radar, *Proceedings of the IEEE,* 51: 1016-1027.

Oppenheim, A. V. and Shafer, R. W. (1989) *Discrete-time signal processing,* Prentice-Hall, Englewood Cliffs, New Jersey.

Oxnard, C. E. (1972) The use of optical data analysis in functional morphology: investigation of vertebral trabecular pattern, in: *The functional and evolutionary biology of primates,* Tuttle, R. (ed)., Aldine-Atherton, Chicago, pp. 337-347.

Oxnard, C. E. (1982) The association between cancellous architecture and loading in bone: an optical data analytic view, *The Physiologist,* 25 (supplement 6): 37-40.

Papoulis, A. (1984) *Probability, random variables and stochastic processes,* 2nd Edition, McGraw-Hill Book Company, New York.

Parfit, A.M. (1979) Quantum concept of bone remodeling and turnover: implications for the pathogenesis of osteoporosis, *Calcified Tissue International,* 28: 1-5

Park, E. A., Jackson, D. and Kajdi, L. (1931) Shadows produced by lead in x-ray pictures of growing skeleton, *American Journal of Diseases of Children,* 41: 485-499.

Pauls, V. and Trott, J. R. (1966) A radiological study of experimentally produced lesions in bone, *Dental Practitioner,* 16: 254-258.

Payton, C. G. (1932) The position of the nutrient foramen and direction of the nutrient canal in the long bones of the madder-fed pig, *Journal of Anatomy,* 68: 500-510.

Phemister, D. B. (1918) Effect of phosporus on growing, normal and diseased bones, *Journal of the American Medical Association,* 70: 1737-1743.

Pocock, S. J. (1983) *Clinical trials: a practical approach,* John Wiley and Sons, New York.

Polig, E. and Jee, W.S.S. (1987) Bone age and remodeling: a mathematical treatise, *Calcified Tissue International,* 41: 130-136

Poswillo, D. (1984) The plain man's guide to research: the design, *British Journal of Oral and Maxillofacial Surgery,* 22: 328-327 (discussion p.328).

Poulton, D. R. (1968) Evaluation of space-closing techniques with the aid of laminagraphic cephalometrics. *American Journal of Orthodontics,* 54: 899-919.

Pratt, W. K. (1991) *Digital image processing,* John Wiley and Sons, New York.

Prewitt, M. S. (1970) Object enhancement and extraction, in: *Picture Processing and Phsychopictorics: Proceedings of the symposium on psychopictorics, Arlington, Virginia, April 1969,* Lipkin B. S. and Rosenfeld, A. (eds)., Academic Press, London, pp. 75-149.

Price, C. (1973) The consistent repetition of periapical radiographs. *M.D.S. Thesis,* Birmingham University Dental School.

Proakis, J. G. and Manolakis, D. G. (1988) *Introduction to Digital Signal Processing,* Macmillan, New York.

Ramadan, A. E. and Mitchell, D. F. (1962) A roentgenographic study of experimental bone destruction, *Oral Surgery, Oral Medicine, Oral Pathology,* 15: 934-943.

Revell, P. A. (1992) Methods of imaging bone (abstract), *Bone,* 13: 95.

Riedel, R. A. (1972) The implant technic including history, relative accuracy and information derived and applied to orthodontic patients, *Bulletin of the Pacific Coast Society of Orthodontists,* 47: 33-42.

Robertson, N. R. E. and Hilton, R. (1971) The changes produced by pre-surgical oral orthopaedics, *British Journal of Plastic Surgery,* 24: 57-68.

Roche, A. F. (1965) Sites of elongation of Human metacarpals and metatarsals, *Acta Anatomica,* 61: 193-202.

Royston, J. P. (1982) An extension of Shapiro and Wilk's W test for normality to large samples, *Applied Statistics,* 31: 115-124.

Rune, B. (1980) Roentgen stereophotogrammetry and metallic implants in the study of craniofacial anomalies, Department of Orthodontics, Faculty of Odontology, University of Lund, Malmo, Sweden.

Rune, B., Selvik, G., Kreigborg, S., Sarnäs, K-V. and Kågström (1979) Motion of bones and volume changes in the neurocranium after craniectomy in Crouzon's disease. A roentgen stereometric study, *Journal of Neurosurgery,* 50: 494-498.

Ruttiman, U. E., Webber, R. L. and Schnidt, E. (1986) A robust digital method for film contrast correction in subtraction radiography, *Journal of Periodontal Research,* 21: 486-495.

Ryan, T. W., Gray, R. T. and Hunt, B. R. (1980) Prediction of correlation errors in stereo-pair images, *Optical Engineering,* 19: 312-322.

Ryan, T. W. and Hunt, B. R. (1981) Recognition of stereo-image cross-correlation errors, in: *Progress in Pattern Recognition,* Volume 1, Kanal, L. N. and Rosenfeld, A. (eds)., North-Holland Publishing Company, New York. pp. 265-322.

Sakrison, D. (1968) *Communication Theory: transmission of waveforms and digital information.* John Wiley and Sons, New York.

Sandor, T., Spears, J. and Paulin, R. (1982) Densitometric determination of changes in the dimensions of coronary arteries. *Society of Photo-optical Instrumentation Engineers: Digital Radiography,* 314: 263-272.

Sarnat, B. G. and Gans, B. J. (1952) Growth of bones: methods of assessing and clinical importance, *Plastic and Reconstructive surgery,* 9: 140-160.

Schmaker, D. B. and El Hadary, M. S. (1960) Roentgenographic study of eruption, *American Journal of Orthodontics,* 61: 535-541.

Scudder, H. J. (1978) Introduction to computer aided tomography, *Proceedings of the IEEE,* 66: 628-637.

Secilla, J. P., Garcia, N. and Carrascosa, J. L. (1988) Template location in noisy pictures, *Signal Processing,* 14: 347-361.

Selvik, G. (1974) A roentgen stereophotogrammetric method for the study of the kinematics of the skeletal system, Thesis, AV-Centralen, University of Lund, Malmo, Sweden.

Shaw, R. (1975) Some fundamental properties of xeroradiographic images, *Society of Photo-Optical Instrumentation Engineers,* 70: 359-363).

Shaw, W. C. (1977) Problems of accuracy and reliability in cephalometric studies with implants in infants with cleft lip and palate, *British Journal of Orthodontics,* 4: 93-100.

Siersbaek-Nielsen, S. (1971) Rate of eruption of central incisors at puberty: an implant study on eight boys, *Tandlaege Bladet,* 75: 1288-1295.

Singh, M., Frei, W., Shibata, T. and Huth G. C. (1979) A digital technique for accurate change detection in nuclear medical images - with application to myocardial perfusion studies using thallium-201, *IEEE Transactions on Nuclear Science,* NS-26: 565-569.

Solow, B. and Kreiborg, S. (1988) A cephalometric unit for research and hospital environments, *European Journal of Orthodontics,* 10: 346-352.

Sontag, L. W. (1938) Evidences of distrubed prenatal and neonatal growth in bones of infants aged one month, *American Journal of Diseases of Children,* 55: 1248-1256.

Steedle, J. R. and Proffit, W. R. (1985) The pattern and control of eruptive tooth movements, *American Journal of Orthodontics,* 87: 56-66.

Stettner, E. (1920) Über die Bezziehungen der Ossifikation des Handskeletts zu Alter und Längenwachstrum bei gesunden und kranken Kindern von der Gebruurt bis zur Puberttät, *Archiv Fur Kinderheilklinik,* 68: 342-368.

Svedlow, M., McGillam, C. D. and Anuta, P. E. (1978) Image Registration: similarity measure and preprocessing method comparisons, *IEEE Transactions on Aerospace and Electronic Systems,* AES-14, 141-149.

Swets, J. A. and Pickett, R. M. (1982) *Evaluation of diagnostic systems: methods from signal detection theory,* Academic Press, New York.

Tan, H. T. and Van Aken, J. (1977) Reliability of orientation planes in facial radiology, in: *Proceedings of the IV International Congress of Dento-Maxillo-Facial Radiology,* Malmö, Sweden, Adatia, A. K., Fujiki, Y., Fuchihata, H., Imai, K., Nakamura, T. and Wada, T. (eds)., International Association of Dento Maxillofacial Radiology, pp. 335-337.

Thunthy, K. H. and Weinberg, R. (1982) Evaluation of contrast in duplicated radiographs, *Oral surgery,* 54: 691-695.

Ton, J. and Jain, A. K. (1989) Registering Landsat images by point matching, *IEEE Transactions on Geoscience and Remote Sensing,* 27: 642-651.

Trouerbach, W. T., Hoornstra, K. and Zwamborn, A. W. (1984) Microdensitometric analysis of interdental bone structure; the development of a registration method, *Dentomaxillofacial Radiology,* 13: 27-31.

Turin, G. L. (1960) An introduction to matched filters, *IEEE Transactions on Information Theory,* IT-6: 311-329.

Van der Stelt, P. F. and Geraets, W. G. M. (1987) Quantitative description of the trabecular pattern of bone in osteoporotic patients, *IEEE Ninth Annual Conference of the Engineering in Medicine and Biology Society,* IEMBS 2: 741-742.

Van Dijk, R., Huiskes, R., and Selvik, G. (1979) Roentgen stereophotogrammetric methods for the evaluation of the three dimensional kinematic behaviour and cruciate ligament length patterns of the human knee joint, *Journal of Biomechanics,* 12: 727-731.

Van Trees, H. (1968) *Detection, Estimation and Modulation theory,* Part I. John Wiley & sons, New York.

Venot, A., Lebruchec J. F., and Roucayrol, J. C. (1984) A new class of similarity measures for robust image registration. *Computer Vision, Graphics and Image Processing,* 28: 176-184.

Voles, R. (1978) Image registration using the likelihood ratio, *Proceedings of the IEE,* 125: 686-690.

Von Neumann, J. and Goldstine, H. (1947) Numerical inverting of matrices of high order, *Bulletin of the American Mathematical Society,* 53: 1021-1099.

Watanabe, M., Uchida, H., Okada, H., Shimizu, K., Sato, N., Yoshikawa, E., Ohmura, T., Yamashita, T. and Tanaka, E. (1992) A high resolution PET for animal studies, *IEEE Transactions on Medical Imaging,* MI-11: 577-580.

Webber, R. L., Ruttiman, U. E. and Groenhuis, R. A. J. (1985) Synthesis of arbitrary projections from a finite number of existing projections, *Proceedings of the Society of Photo-Optical Instrumentation Engineers: Applications of Optical Instrumentation in Medicine XIII,* 535: 84-91.

Whalen, A. D. (1971) *Detection of signals in noise,* Academic Press, New York.

Wiles, C. and Foreshaw, M. R. (1993) Recognition of volcanoes on Venus using correlation methods, *Image and Vision Computing,* 11: 188-196.

Woods, J. W. (1972) Two dimensional discrete Markov fields, *IEEE Transactions on Information Theory,* IT-18: 232-240.

APPENDIX A

FORMULATION OF THE DECORRELATOR OR "WHITENING FILTER"

A.1 PREAMBLE

To allow the detection processor to operate correctly the correlated temporal noise must be converted to an equivalent uncorrelated form. As explained in Chapter four, this conversion is achieved by passing the temporal noise through an initial processor - the decorrelator or "whitening filter". The first-order MRF model proposed for the temporal noise allows the systematic development of a decorrelator based on the spatial ACF of the model. This section describes the formulation of the decorrelator in the spatial domain, that is, its impulse response.

A.2 FORMULATION OF THE DECORRELATOR

This method is based on the solution of the discrete matched filter equation for two-dimensional signals derived by Arcese, Mengert and Trobini (1970); and by Prewitt (1970). The reason for using this approach is that the discrete matched filter equation defines the operations necessary to decorrelate an image and thereby achieve the maximum signal-to-noise ratio in a correlation detector.

To simplify the analysis two additional assumptions have been made about the MRF model of the temporal noise image. These are: first, that the ACF of the image model is of "separable" form; and second, that the correlation between contiguous pixels (the "one-step correlation") is the same in both axes of the image.

The determination of the impulse response requires the numerical solution of the discrete matched filter equation, which, in matrix form is:

$$\mathbf{\Sigma\Psi} = \mathbf{\Phi} \qquad \text{(Eq. A.1)}$$

where $\mathbf{\Phi}$ is the matrix representation of the (correlated) temporal noise image; $\mathbf{\Psi}$ is the matrix representation of the uncorrelated temporal noise image; and $\mathbf{\Sigma}$ is the covariance matrix of the temporal noise.

To represent the image as a suitable matrix it is first necessary to scan the image in horizontal lines to produce a row (or column) of numbers representing the pixel values in each line of the image. Each row of numbers is known as a row vector. All the row vectors are then placed together in a larger matrix representing the whole image. The matrix formed from these row vectors is known as the image matrix and is symbolised by $\mathbf{\Phi}$ in equation A.1. A second, similar matrix whose elements are the row vectors of the decorrelated temporal noise image is symbolised by $\mathbf{\Psi}$. The final matrix in equation A.1, known as the covariance matrix, $\mathbf{\Sigma}$, consists of elements, Σ_{ij}, which are the cross-correlation matrices of the image rows.

The solution to equation A.1 is simply

$$\mathbf{\Psi} = \Sigma^{-1}\mathbf{\Phi} \qquad \text{(Eq. A.2)}$$

However, to solve the equation numerically requires the inversion of the covariance matrix. Not all matrices can be explicitly inverted and even where inversion is possible the process is computationally prohibitive for all but the largest computers (Pratt, 1978). However, images which can be modelled as separable first-order MRF's offer two computational advantages which simplify the calculation: first, the covariance matrix of an MRF has a specific form known as "block Toeplitz" which can be inverted; and Secondly, the cross-correlation matrices, Σ_{ij}, (the elements of the covariance matrix) have a well defined (and analytically tractable) exponential structure related to the correlation, ρ, between the pixels in the MRF image. Thus: the cross-correlation matrix between the *i*th and *j*th rows of the image will be,

$$\Sigma_{ij} = \rho^{|i-j|} \Sigma_{ii} \tag{Eq. A.3}$$

Where, Σ_{ii}, is the autocorrelation matrix of the *i*th row. Thus, because Σ_{ii}, is a constant, Σ_{ij} will decline as a power of ρ (that is, exponentially) with distance between the *i*th and *j*th rows. Similarly, autocorrelation matrix of the *i*th row of the image, Σ_{ii} will have the exponential form,

$$\Sigma_{ii} = (\rho^{|i-j|}) \tag{Eq. A.4}$$

Consequently, the covariance matrix reduces to a block Toeplitz matrix whose elements are simply integral powers of the one-step correlation, ρ. The inverse of the image covariance matrix, Σ^{-1} is given in Figure A1.

To decorrelate the image it is necessary to multiply the image vector by this matrix-inverse. However, the decorrelator must be implemented as a linear processor. In order to develop a decorrelator which is also a linear system, the decorrelation must be carried out as a convolution operation on the original image rather than using image matrices and matrix multiplication. Accordingly, a new function must be defined which is the equivalent of the covariance matrix-inverse but which can be convolved directly with the temporal noise image to produce the required decorrelation. The form of this function can be determined by considering the actual components of the individual vectors which form the decorrelated image matrix, Ψ.

The *k*th component of the *i*th vector $\Psi(i)$, Ψ_{ik}, can be represented in term of its components, as

$$\Psi_{ik} = -\rho[-\rho\Phi_{i-1,k-1} + (1+\rho^2)\Phi_{i-1,k} - \rho\Phi_{i-1,k+1}] + (1+\rho^2)[-\rho\Phi_{i,k-1} + (1+\rho^2)\Phi_{i,k} - \rho\Phi_{i,k+1}] - \rho[-\rho\Phi_{i+1,k-1} + (1+\rho^2)\Phi_{i+1,k} - \rho\Phi_{i+1,k+1}] \tag{Eq. A.5}$$

Where each element, Φ, of the image row vector represents a specific pixel in the uncorrelated image; and the subscripts associated with each element indicate the

---

$$\Sigma^{-1} = 1/(1-\rho^2)\begin{bmatrix} 1 & -\rho\Sigma_{ii}^{-1} & 0 & \cdots & \cdots & 0 \\ -\rho\Sigma_{ii}^{-1} & (1+\rho^2)\Sigma_{ii}^{-1} & -\rho\Sigma_{ii}^{-1} & 0 & \cdots & 0 \\ 0 & -\rho\Sigma_{ii}^{-1} & (1+\rho^2)\Sigma_{ii}^{-1} & -\rho\Sigma_{ii}^{-1} & 0 & 0 \\ \vdots & 0 & & & & \vdots \\ 0 & \cdots & \cdots & 0 & -\rho\Sigma_{ii}^{-1} & 1 \end{bmatrix}$$

$$\Sigma_{ii}^{-1} = 1/(1-\rho^2)\begin{bmatrix} 1 & -\rho & 0 & 0 & \cdots & . & 0 \\ -\rho & (1-\rho^2) & -\rho & 0 & \cdots & . & 0 \\ 0 & -\rho & (1-\rho^2) & -\rho & \cdots & . & 0 \\ \cdot & & & & & & \cdot \\ 0 & \cdots & . & . & 0 & -\rho & 1 \end{bmatrix}$$

where Σ_{ii}^{-1} is the inverse of the autocorrelation matrix of the temporal noise image; ρ ($= \rho_m = \rho_n$) is the one-step correlation for the MRF temporal noise image model.

Figure A1 Covariance matrix inverse for the first-order MRF model.

---

spatial position of the pixel in the image. By expanding the left-hand side of equation A.5 and ordering the resulting terms into lines which correspond with the relative spatial positions of the pixels in the decorrelated image,

$$\begin{aligned}\Psi_{ik} = \quad & \rho^2\Phi_{i-1,k-1} - \rho(1+\rho^2)\Phi_{i-1,k} + \rho^2\Phi_{i-1,k+1} \\ & - \rho(1+\rho^2)\Phi_{i,k-1} + (1+\rho^2)^2\Phi_{i,k} - \rho(1+\rho^2)\Phi_{i,k+1} \\ & \rho^2\Phi_{i+1,k-1} - \rho(1+\rho^2)\Phi_{i+1,k} + \rho^2\Phi_{i+1,k+1}\end{aligned} \qquad \text{(Eq. A.6)}$$

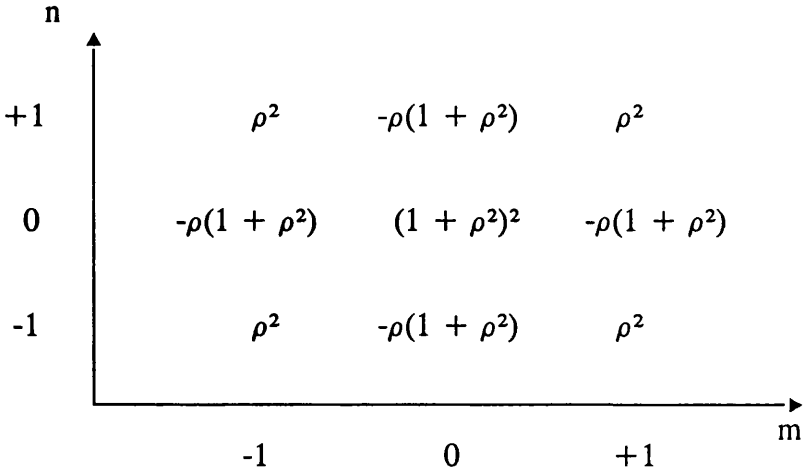
the effect of the decorrelation can be seen in terms of the components of the original correlated image. Thus, the three lines of equation A.6 represent three vertically adjacent horizontal lines of three pixels each. Therefore, a block of 3 x 3 pixels is involved at the *k*th component of the *i*th vector $\Psi(i)$, Ψ_{ik}. Accordingly, to construct the *ik*th component of the decorrelated image vector using a convolution operation, involves multiplying the 3 x 3 block of pixels from the original correlated image by a two-dimensional function, which we shall refer to as W_i, followed by summation of the resulting terms.

The form of the function, W_i, can be deduced from the coefficients of the row vector elements ($\Phi_{i,k}$). Referring again to equation A.6, the first line of the equation indicates that the three pixels from left to right have the coefficients: ρ^2; $-\rho(1-\rho^2)$; and $-\rho^2$ respectively. For the second line the coefficients are: $-\rho(1+\rho^2)$; $(1+\rho^2)^2$ and $-\rho(1+\rho^2)$ respectively. The coefficients of the third line repeat those in the first line. Therefore, to execute the decorrelation as a convolution operation (performed by a linear processor) each block of 3 x 3 pixels in the image is multiplied by the two-dimensional function,

$$W_i = \left\{ \begin{array}{ccc} \rho^2 & -\rho(1+\rho^2) & \rho^2 \\ -\rho(1+\rho^2) & (1+\rho^2)^2 & -\rho(1+\rho^2) \\ \rho^2 & -\rho(1+\rho^2) & \rho^2 \end{array} \right. \qquad \text{(Eq. A.8)}$$

followed by summation of the resulting terms.

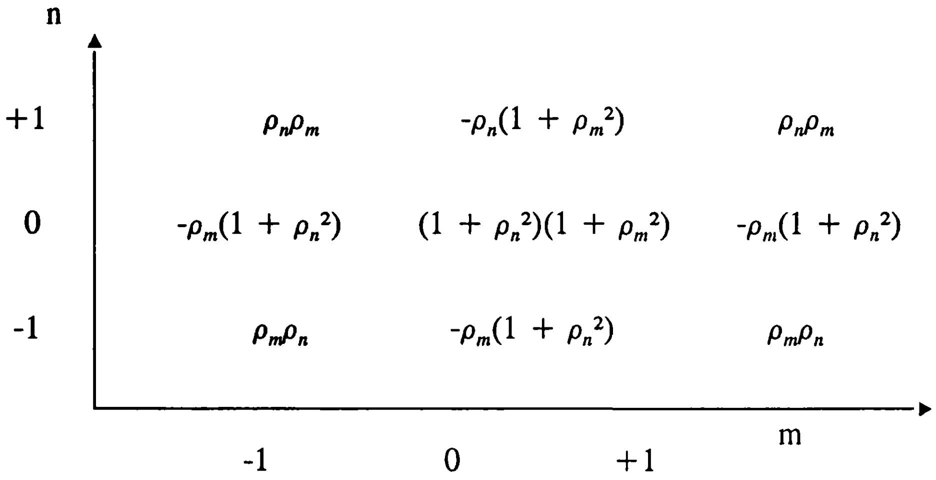
Where the one-step correlation of the MRF model is the same in both planes of space, the impulse response of the decorrelator may be represented by a spatial representation of the function W_i shown in Figure A2.



n \ m	-1	0	+1
+1	ρ^2	$-\rho(1 + \rho^2)$	ρ^2
0	$-\rho(1 + \rho^2)$	$(1 + \rho^2)^2$	$-\rho(1 + \rho^2)$
-1	ρ^2	$-\rho(1 + \rho^2)$	ρ^2

Figure A2 The impulse response for the decorrelator where ρ ($= \rho_m = \rho_n$) is the one-step correlation for the MRF temporal noise image model.

However, if the one-step correlation is different in the horizontal and vertical directions, the impulse response of the decorrelator has a similar form but the coefficients are factored into horizontal and vertical components (Figure A3).



n \ m	-1	0	+1
+1	$\rho_n\rho_m$	$-\rho_n(1 + \rho_m^2)$	$\rho_n\rho_m$
0	$-\rho_m(1 + \rho_n^2)$	$(1 + \rho_n^2)(1 + \rho_m^2)$	$-\rho_m(1 + \rho_n^2)$
-1	$\rho_m\rho_n$	$-\rho_m(1 + \rho_n^2)$	$\rho_m\rho_n$

Figure A3 The impulse response for the decorrelator where the one-step correlations for the MRF temporal noise image model differ in the two planes of space ($\rho_m \neq \rho_n$).

APPENDIX
B

VERIFICATION OF IMPLANT STABILITY

B.1 Procedure

The initial selection of the radiographs in data sets 1 and 2 (Figure B1) was made on the basis that the implant images could be aligned by direct superimposition of the films. While this provided a guide to their likely stability it was necessary to establish an objective comparison of their inter-relationships and thereby confirm their stability, within the limits of the two-dimensional image.

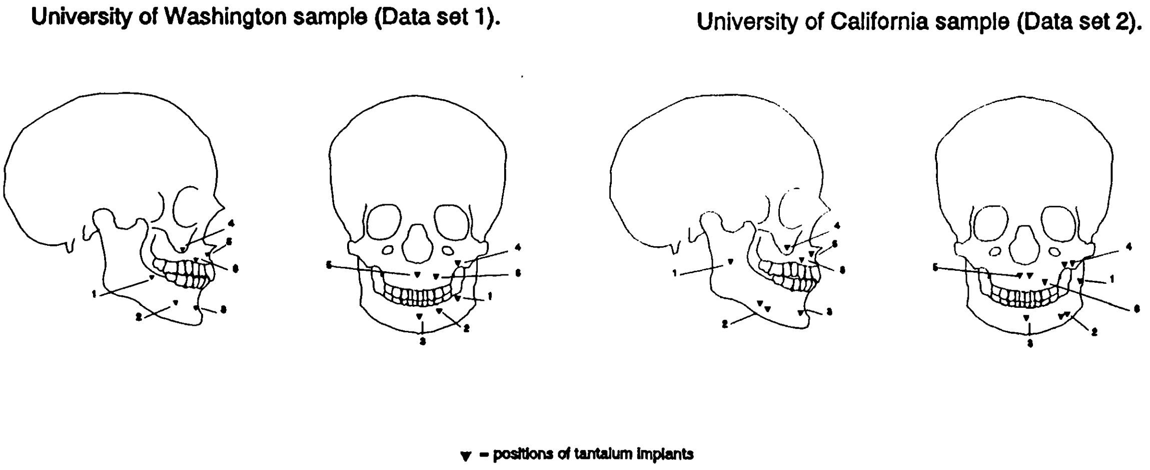


Figure B1 The location of the tantalum implants in data sets 1 and 2.

The inter-relationships between the implants was measured using an electronic digitizing table(16). Each duplicated radiograph was secured in turn to the illuminated digitizing table and the centre of each implant was located and digitized using the hand held cursor. The distances between pairs of implants, within the same jaw, were then recorded and compared between pairs of duplicated radiographs.

The criterion for inclusion in the final data sets was that the maximum absolute difference in inter-implant distances should be no greater than 0.4 mm for at least one pair of implants in each jaw. Implant pairs not fulfilling this criterion were excluded from the study.

B.2 Results

The absolute difference in inter-implant distance are shown in Table B1. It is important to note that for each subject each implant was included in two measurements for each set of three implants (Figure B2). Fourteen implant pairs were rejected in the mandible and twelve in the maxilla of data set 1; 2 implant pairs were rejected in both the mandible and maxilla of data set 2.

The error of measurement was determined from 20 duplicate measurements of the same inter-implant distances made on two separate occasions. The mean absolute difference between duplicate measurements was 0.12 mm.

Table B1 Absolute differences (mm) in paired inter-implant distances.

		Mandible		Maxilla	
	n	mean	maximum	mean	maximum
Data set 1	150	0.17	0.40	0.13	0.40
Data set 2	12	0.19	0.25	0.15	0.37

Notes: n = number of paired measurements;

50 cases in data set 1; 4 cases in data set 2.

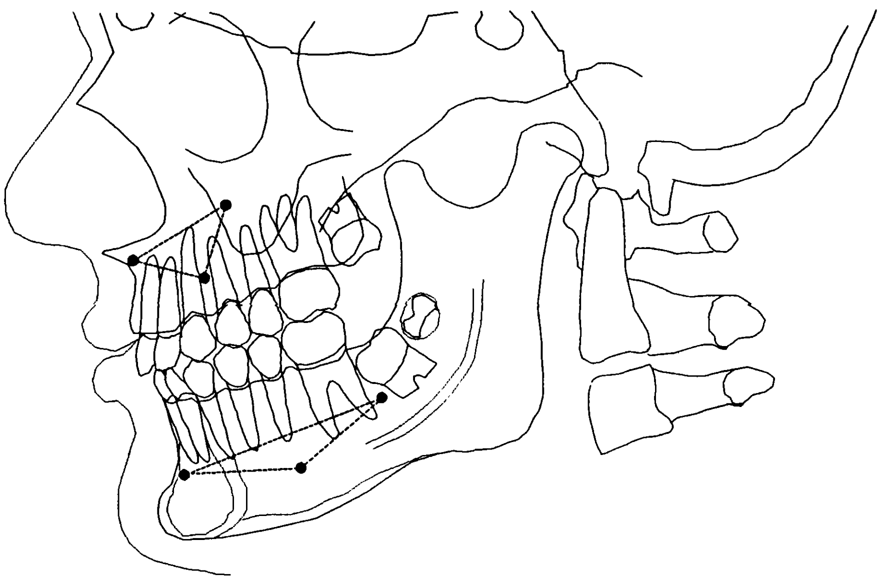


Figure B2 Inter-implant distances measured to determine the level of implant stability: ------ represent the inter-implant lines whose distances were compared between pairs of films.

B.3 Conclusions

Within the limits imposed by two-dimensional images, the radiographs employed in this study depicted implants whose instability was no greater than 0.4 mm for all implant pairs.

APPENDIX C

MANUFACTURERS AND SUPPLIERS OF MATERIALS

(1) Kodak X-Omat duplicating film: Kodak Ltd., Hemel Hempstead, Herts., UK.

(2) RCA TC 1005 Vidicon Camera: Professional Products (UK), Alperton House, Wembley, HA0 1EG., UK.

(3) National Panasonic WV-5310, Panasonic House, Bracknell, Berkshire, UK.

(4) Computer Recognition Systems Ltd, Wokingham, Berkshire, UK.

(5) Sun Microsystems Europe Inc., Camberley, Surrey, UK.

(6) Apple Computer Inc, Cupertino, California, USA

(7) Bell Laboratories, New Jersey, USA.

(8) Graticules Ltd., Tonbridge, Kent, UK.

(9) Picture Processing Software, Human Information Processing Laboratory, Psychology Dept, New York University, New York, USA.

(10) SPSS Inc. 444 N. Michigan Avenue, Chicago, USA.

(11) Provalis Research, Montreal, Canada.

(12) Medical Computing, Augthon, Lancashire, UK.

(13) CRC Graphics Package; Sun Microsystems Europe Inc., Camberley, Surrey, UK.

(14) IBM UK Ltd., London W4 5TF.

(15) Viglen Ltd., Alperton, Middlesex, UK.

(16) Lysta Dent electro-optical digitizing table, Lysta Ltd., 92 Church Road, Mitcham CR4 3TD, UK.

APPENDIX
D

ESTIMATION OF THE ONE-STEP CORRELATION ($\rho_{one\text{-}step}$) OF THE TEMPORAL NOISE

This appendix describes the experimental estimation of the one-step correlation of the temporal noise derived from a sample of serial radiographs registered by reference to implanted tantalum markers.

D.1 Materials and method

The one-step correlations in the horizontal and vertical planes of the radiographic images were estimated from pairs of serial lateral cephalometric radiographs of the mandible and maxillae derived from the twenty-four subjects in data set 1a.

As in the validation study, images from the earlier and later radiographs of each pair were recorded and stored as a series of subimages (approximately 1.55 cm^2) of 96 x 96 pixels with a pixel size of 0.13 mm x 0.13 mm from pre-determined sites. For each subject a single subimage was selected from each jaw. The choice of subimage was made randomly from the sites in each jaw as described in section 7.5.8 (see Figure 7.5). The subimages were then registered by using the tantalum markers as described in section 7.5.7 and the earlier images subtracted from the later image and the difference image stored.

A single 40 x 40 pixel samples was drawn from the mandibular and maxillary subimages of each difference image. The spatial positions of the samples were determined randomly within each the central 90 x 90 pixels of each subimage. The one-step correlations were calculated for each 40 x 40 pixel sample located using the computer program **SIMPLE_CORREL** and the resulting normalised SSD correlation function converted to a correlation coefficient using the formula given in Appendix E.2.

D.2 Results

The results are presented in Table D1.

Table D1 Estimated Horizontal one-step correlation coefficient of the temporal noise.

subject number	Mandible one-step correlation coefficient ($\rho_{1\text{-step}}$) x	Mandible y	Maxilla one-step correlation coefficient ($\rho_{1\text{-step}}$) x	Maxilla y
1	0.63	0.65	0.62	0.65
2	0.71	0.58	0.61	0.70
3	0.69	0.37	0.47	0.51
4	0.45	0.61	0.52	0.67
5	0.53	0.62	0.55	0.48
6	0.58	0.67	0.46	0.57
7	0.79	0.58	0.55	0.46
8	0.64	0.65	0.52	0.56
9	0.56	0.54	0.61	0.69
10	0.63	0.30	0.48	0.51
11	0.45	0.48	0.52	0.58
12	0.50	0.62	0.55	0.49
13	0.58	0.62	0.49	0.56
14	0.79	0.58	0.56	0.46
15	0.60	0.63	0.52	0.65
16	0.36	0.58	0.63	0.71
17	0.69	0.35	0.70	0.51
18	0.48	0.61	0.39	0.52
19	0.55	0.62	0.55	0.46
20	0.58	0.67	0.47	0.50
21	0.60	0.59	0.60	0.46
22	0.65	0.65	0.41	0.56
23	0.49	0.54	0.50	0.64
24	0.44	0.30	0.44	0.52

The distribution of values for each coordinate axis were examined for normality using the Shapiro-Wilk test. In no case was the distribution significantly different from normal. The $\rho_{1\text{-step}}$ values were compared between the coordinate axes and between the two jaws using a two-way analysis of variance. No significant difference was found for the $\rho_{1\text{-step}}$ values between the coordinate axes within or between the two jaws (variance ratio between jaws, F=1.03; p=0.44; variance ratio between axes, F=1.19; p=0.32).

Using the combined data, the mean $\rho_{1\text{-step}}$ was 0.56 (SD 0.095).

D.3 Conclusions

The lack of clear evidence of a difference in the $\rho_{1\text{-step}}$ values between the coordinate axes and between the jaws supports the idea of circular symmetry for the temporal noise ACF. However, the standard deviation of the $\rho_{1\text{-step}}$ values, which was close to 0.1, makes it difficult to set a single value for the $\rho_{1\text{-step}}$ which is appropriate for all samples of noise. However, because the impulse response of the decorrelator is restricted to integer values in a 3 x 3 array, the practical effect of choosing a single mean value is likely to be limited. Consequently, the $\rho_{1\text{-step}}$ was set just below the mean at 0.50 for both coordinate axes of the mandible and maxilla to avoid over decorrelation (forcing the internal correlation of decorrelated signal and noise below zero) in those samples with the lowest $\rho_{1\text{-step}}$ values.

Having set the $\rho_{1\text{-step}}$ value at 0.50 the kernal values for the decorrelator impulse response array were determined from equation A.8, rounded and converted to integer values to provide the impulse response array:

1	-3	1
-3	6	-3
1	-3	1

APPENDIX
E

NORMALISED SSD CORRELATION FUNCTION

E.1 Normalised SSD correlation function

The normalised SSD correlation function, *C(p,q)*, is expressed mathematically as:

$$C(p,q) = \sum_{i,j=0}^{N-1} \left\{ \left(\frac{t(i,j) - \bar{t}}{\sqrt{\sum_{i,j=0}^{N-1} [t(i,j) - \bar{t}]^2}} \right) - \left(\frac{I(i+p, j+q) - \bar{I}}{\sqrt{\sum_{m,n=0}^{N-1} [I(m+p, n+q) - \bar{I}]^2}} \right) \right\}^2$$

(Eq. D.1)

where:

The template comprises $N \times N$ pixels and;
I represents the pixel values of the image (search) area;
$\bar{I}$ is the local mean image value;
t represents the pixel values of the template;
$\bar{t}$ is the mean template value;
i; j; k; l; m; n; p; and *q* are indices which represent spatial coordinates;

E.2 Conversion of normalised SSD values to a correlation coefficient (ρ)

The normalised SSD correlation function is converted to a correlation coefficient using the following formula:

$$\rho = 1 - \mathrm{SSD}^{1/2} \qquad \text{(Eq. E.2)}$$

where:

ρ is the correlation coefficient;

$\mathrm{SSD}^{1/2}$ is the positive square root of the normalised SSD correlation.

APPENDIX F

EXPERIMENT TO ESTABLISH THE DETECTION THRESHOLD

The probability of a correct match is determined by the SNR_{out}, which is defined as the ratio of the correlation at the peak to the standard deviation of the correlations of the template with the surrounding background (Ryan et al., 1981).

To reduce the probability of falsely detecting a stable site during stage 2 of the detection system, a threshold was set for the acceptance of the peak as indicating a true match. The threshold was chosen such that only rarely does a peak value obtained from an incorrect match reach above this threshold, giving rise to a false match.

The threshold can be determined from theoretical considerations if the probability density functions of correct and incorrect outputs are known or assumed. However, in the absence of previous studies on detection of structures in bone, the threshold was determined empirically. This is done by measuring the probability density function of the correlations between the template and the regions surrounding the true peak.

F.1 Materials and methods

The background PDF was estimated from the correlation values obtained between non-matching sites in the mandible and maxillae of subjects in data set 1a. To provide a large enough sample to estimate the PDF of false matches it was necessary to consider multiple observations from the same subject as being statistically independent. Nevertheless, to avoid gross dependence between observations, samples from the same individual were recorded from non-overlapping sites, but within the same bone.

F.1.1 *Image capture and preliminary processing.*

Images from the earlier and later radiographs of each pair were captured as a series of subimages of 96 x 96 pixels with a pixel size of 0.13 mm from two anatomical sites in the mandible and two sites in the maxillae (Figure F1) as described in section 7.5. Corresponding subimages from each pair of radiographs were registered using the tantalum markers as described in section 7.5.8.

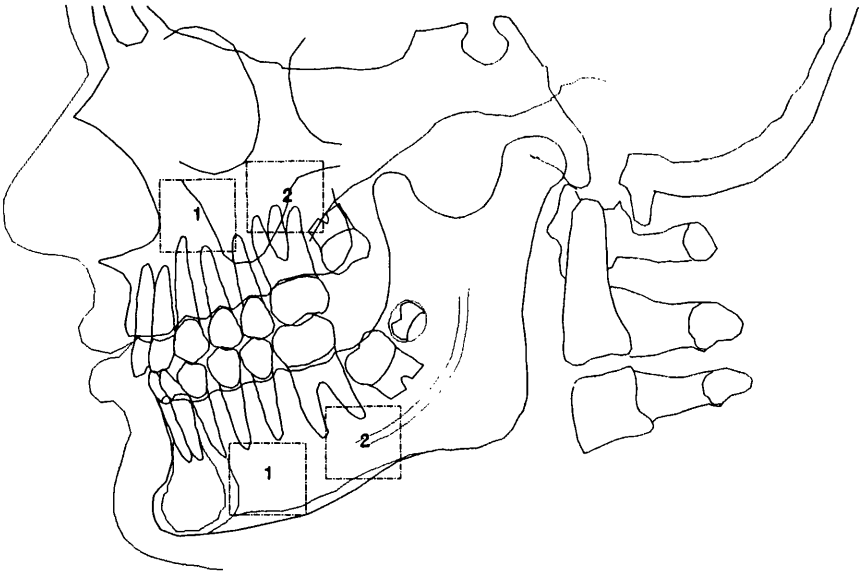


Figure F1 Diagram showing sites of subimages.

The subimages from the later radiograph were convolved with a 3 x 3 Sobel gradient impulse response array using the YAM convolve function. The resultant gradient image was contrast stretched using the YAM enhance function and thresholded at the median grey level value. Those pixels below the threshold were set to zero; those above the threshold to unity. The resulting binary image was then processed to locate square regions 15 x 15 pixels containing high variance data. This was done by convolving the binary image with a unit valued image of 15 x 15 pixels. The coordinates of the region with the highest proportion of high variance data were recorded and used to position the template in the original (unprocessed) later image.

The template from the later image was then cross-correlated with four 15 x 15 pixel sections from the non-corresponding subimage of the earlier image, using the program CORREL. The peak cross-correlation values obtained with the non-corresponding section were used to estimate the cross-correlation PDF for false-positive detection.

F.2 Results

Results are shown in Table F1. Prior to setting the threshold the distributions were tested for conformity to a normal distribution using the Shapiro-Wilk W test.

Table F1 Distribution of cross-correlation coefficients for false-positive detection.

		Shapiro-Wilk test				
	n	W	p	mean	σ	2.5σ above mean
Mandible	96	0.947	0.211	56.73	4.54	68.08
Maxilla	96	0.912	0.237	59.47	6.34	75.32

F.3 Conclusions

Using the estimates obtained in this experiment, the threshold was set at a level which corresponds to 2.5 standard deviations (2.5σ) above the mean of the background correlations which provides a fixed false-positive probability of approximately $p=0.005$. Therefore the thresholds were set at $\rho=0.68$ and $\rho=0.75$ for the mandible and maxilla respectively.

APPENDIX G

SOURCE CODES FOR COMPUTER PROGRAMS

G.1 IMAGE CAPTURE AND NOISE FILTERING PROGRAM

G.1.1 YAM program file "GRAB" - Professor M. Day modified by S. D. Springate

This program grabs and averages 48 individual images from each of two radiographs. The images are acquired from the video camera via the framestore in block of 8 frames. On completion of the averaging, the two images are stored in temporary memory as tempory storage files *STO0* and *STO6*.

The program is interactive requiring actions indicated by the display on the computer screen.

```
LAST=PRIMARY;
echo("align first image, press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO1=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO2=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO3=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
```

```
wf(PRIMARY,".yamtool.tmpfile");
STO4=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO5=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO7=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO8=PRIMARY;
LAST=PRIMARY;
PRIMARY=(STO1+STO2+STO3+STO4+STO5+STO7+STO8)/6; display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO0=PRIMARY;
LAST=PRIMARY;
echo("align second image, press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO1=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO2=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO3=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO4=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
```

```
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO5=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO7=PRIMARY;
LAST=PRIMARY;
echo("press <esc> when the image is acceptable")
while notzero(1) di(grab(8)) endwhile grab8=0 do 8
grab8=grab8+grab(8) enddo PRIMARY;
wf(PRIMARY,".yamtool.tmpfile");
STO8=PRIMARY;
LAST=PRIMARY;
PRIMARY=(STO1+STO2+STO3+STO4+STO5+STO7+STO8)/6; display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO6=PRIMARY;
LAST=PRIMARY;
```

G.2 DECORRELATION PROGRAM

G.2.1 YAM program file "DECORREL" - Author S. D. Springate

This program will Gaussian noise-filter, decorrelate, contrast-stretch two images (located in image files *image,0* and *image,1)* stored in central memory. On completion of the decorrelation the processed images are stored in central memory as *whiteimage,0* and *whiteimage,1* (additional copies are made and stored in temporary memory).

The program is interactive requiring the image file names (*image,0* and *image,1*) to be typed and entered using the readfile key.

```
PRIMARY=rf("image,0");
display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO0=PRIMARY;
LAST=PRIMARY;
PRIMARY=rf("image,1");
display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO6=PRIMARY;
LAST=PRIMARY;
PRIMARY=rf("ONEMASK",0);
```

```
display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO7=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO0;
display(PRIMARY);
#include "gauss"
wf(PRIMARY,".yamtool.tmpfile");
STO0=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO6;
display(PRIMARY);
#include "gauss"
wf(PRIMARY,".yamtool.tmpfile");
STO6=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO0;
display (PRIMARY);
#include "stat35";
#include "terry";
wf(PRIMARY,".yamtool.tmpfile");
STO0=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO6;
display (PRIMARY);
#include "stat35";
#include "terry";
wf(PRIMARY,".yamtool.tmpfile");
STO6=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO7;
display (PRIMARY);
LAST=PRIMARY;
SECONDARY=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO0;
display (PRIMARY);
LAST=SECONDARY;
PRIMARY*=SECONDARY;
display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO0=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO7; display (PRIMARY);
LAST=PRIMARY;
SECONDARY=PRIMARY;
LAST=PRIMARY;
PRIMARY=STO6;
display (PRIMARY);
LAST=SECONDARY;
PRIMARY*=SECONDARY;
```

```
display(PRIMARY);
wf(PRIMARY,".yamtool.tmpfile");
STO6=PRIMARY;
LAST=PRIMARY;
wf(PRIMARY,"whiteimage,1")
LAST=PRIMARY;
PRIMARY=STO0; display (PRIMARY);
wf(PRIMARY,"whiteimage,0")
```

G.2.2 Secondary (sub)programs and YAMfiles activated by DECORREL

G.2.2.1 *YAM program file "Stat35" - Author S. D. Springate*

This program was designed to provide a convolution between the image and a 3 x 3 integer valued impulse response array.

```
PRIMARY=convolve(PRIMARY,1,-3,1,-3,6,-3,1,-3,1);
PRIMARY=(PRIMARY)/4;
di(PRIMARY);
```

G.2.2.2 *YAM program file "terry" - Author Dr. T. Fountain*

This program was designed to adjust the image contrast to fill the full (8 bit) range. The code used in "terry" is given below. The code for *enhance.yam* is not available. However, *enhance.yam* is a standard image contrast enhancement program similar to those found in most image processing packages.

```
#include "/home/ipga/terry/YAM/progs/enhance.yam"
```

G.2.2.3 *YAM program file "Gauss" - Author S. D. Springate*

This program is designed to provide a 3 x 3 convolution between the image and a 3 x 3 integer valued Gaussian impulse response array.

```
PRIMARY=convolve(PRIMARY,1,2,1,2,4,2,1,2,1);
PRIMARY=(PRIMARY)/16;
di(PRIMARY);
```

G.2.2.3 *YAMfile "ONEMASK" - Author S. D. Springate*

This file contains a binary valued image of 96 x 96 pixels, known as a "mask". The central 94 x 94 pixels have a value of unity the remainder have a value of zero.

G.3 **CROSS-CORRELATION PROGRAMS**

G.3.1 **C program file "CORR" - Author Dr. C. Wiles**

```
corr.c        Fri Jun 12 09:31:39 1992        1


/*
 * corr.c:  auto or cross-correlate a template over an image
 *
 * Description:
 *
 * Program to read an NxN (96x96) image into an array, then either:
 *         extract a section from the image to use as a template;
 *         or load in a template from a file.
 * Next auto or cross-correlate the template across the whole image.
 *
 *
 * Compilation:
 *         cc -O2 -o corr corr.c -lm
 *
 */
#include <stdio.h>
#include <string.h>
#include <math.h>


#define TRUE     1
#define FALSE    0

#define OUTPUT_EXTRA_INFO  TRUE    /*output some debug/performance data*/

#define N                 96    /* Size of input/output images (in pixels) */

#define MAX_T_SIZE        25    /* Arbitrary upper limit on template sizes */
#define T_SIZE_DEF        15    /* Default size of template in pixels (MUST BE ODD)*/

#define X_START_DEF       48    /* Default location of the centre of the */
#define Y_START_DEF       48    /* template to be correlated (arbitrary) */

#define EDGE_VAL         255    /* Value for pixels at edge of correlation map */

char INPUT_TEMPLATE = FALSE, OUTPUT_TEMPLATE = FALSE;     /*Template i/o flags*/
char SET_COORDS = FALSE;                               /*Template coordinate setting flag*/
```

```
main(argc, argv)
int argc;
char **argv;
{
char *command = argv[0];       /*Store command name*/
char *argument;                /*Variable to hold current command line argument*/
char *infile = NULL, *outfile = NULL;     /*Define i/o filename strings          */
char *mapfile = NULL;
char *templateout, *templatein;     /*Declare template i/o filename strings    */
FILE *f_out, *f_in, *f_map,    /*Declare FILE pointers for image i/o          */
     *ftem_out, *ftem_in;      /*  and for template i/o                       */
void usage();           /*Function to print "usage" message on error    */
unsigned char image[N][N];                      /*Array to hold image         */
int template[MAX_T_SIZE][MAX_T_SIZE];           /*Array to hold template      */
double norm_templ[MAX_T_SIZE][MAX_T_SIZE]; /*Array to hold normalised template*/
int templ_dc, local_img_dc;      /*Variables to hold mean (dc) values         */
int imval_dc;
double norm_imval;
int templ_ssd, local_img_ssd;
double templ_rssd, local_img_rssd;
int x, y;                        /*Image coordinate variables                 */
int local_x, local_y;            /*Template coordinate variables              */
int t_size = T_SIZE_DEF;         /*Set template size to default value         */
int t_xstart = X_START_DEF;    /*Set template location coords to default values*/
int t_ystart = Y_START_DEF;    /*(not used if template is input from a file)  */
double corr_val, diff;        /*Correlation value and image/template difference*/
double corr_map[N][N];        /*Array to hold correlation map                 */
double max_corr_val;             /*Variable to hold max correlation value     */
double min_corr_val = 999999.9;
double scale_factor;    /*Variable used to scale correlation values from 0-255*/
               /*(this allows the correlation values to be output as an image)*/
double sum_corr_val, mean_corr_val, sd_corr_map;
double threshold, nsds;
int count;

/*
 *  Print out usage information if user requests help
 */
  if( argc == 2 && (!strcmp(argv[1],"-help") || !strcmp(argv[1],"=")) )
  {
    usage(command);
    exit(0);
  }

/*
 *  Set the correlation map stats variables to zero
 */
  sd_corr_map = sum_corr_val = max_corr_val = (double)(count = 0);

/*
 *  Parse command line arguments
 */
  while( --argc )
  {
    argument = *++argv;

    if( !strncmp("-o", argument, 2) )          /*Output thresholded correlation*/
    {                                          /*map file*/
      if( strlen(argument) > 2)
        outfile = argument + 2;
      else if( --argc <= 0)
      {
        usage(command);
        exit(-1);
      }
```

```
      else
        outfile = *++argv;
    }

    else if( !strncmp("-m", argument, 2) )  /*Output full correlation map file*/
    {
      if( strlen(argument) > 2)
        mapfile = argument + 2;
      else if( --argc <= 0)
      {
        usage(command);
        exit(-1);
      }
      else
        mapfile = *++argv;
    }

    else if( !strncmp("-tout", argument, 5) )       /*Template storage file*/
    {
      OUTPUT_TEMPLATE = TRUE;
      if( strlen(argument) > 5)
        templateout = argument + 5;
      else if( --argc <= 0)
      {
        usage(command);
        exit(-1);
      }
      else
        templateout = *++argv;
    }

    else if( !strncmp("-tin", argument, 4) )         /*Input template file*/
    {
      if( SET_COORDS )                   /*Template coordinates have been set*/
      {                                  /*so specifying template input file */
        usage(command);                              /*is pointless!   */
        fprintf(stderr, "('-c' and '-tin' options are mutually exclusive)\n");
        exit(-2);
      }
      INPUT_TEMPLATE = TRUE;
      if( strlen(argument) > 4)
        templatein = argument + 4;
      else if( --argc <= 0)
      {
        usage(command);
        exit(-1);
      }
      else
        templatein = *++argv;
    }

    else if( !strncmp("-s", argument, 2) )       /*Specified template size*/
    {
      if( strlen(argument) > 2)
        t_size = atoi(argument+2);
      else if( --argc <= 0)
      {
        usage(command);
        exit(-1);
      }
      else
        t_size = atoi(*++argv);
    }
```

```
    else if( !strcmp("-c", argument) )    /*Template coordinates within image*/
    {
      if( INPUT_TEMPLATE )         /*Template already being input from a file*/
      {                            /*so specifying template coordinates is   */
        usage(command);                                     /* pointless!  */
        fprintf(stderr, "('-c' and '-tin' options are mutually exclusive)\n");
        exit(-2);
      }
      else if( --argc,--argc <= 0)              /*Two arguments must follow*/
      {
        usage(command);
        exit(-1);
      }

      SET_COORDS = TRUE;          /*Flag the setting of the template coords*/
      t_xstart = atoi(*++argv);  /*NB. These are coords of the template-centre*/
      t_ystart = atoi(*++argv);  /*adjusted later to template top-left corner*/
    }

    else if( infile == NULL )
      infile = argument;

    else
    {
      usage(command);
        exit(-1);
    }
  }

/*
 *  Check template size is legal and odd
 */
  if(t_size < 3 || t_size > MAX_T_SIZE)
  {
    fprintf(stderr,"%s: template size must be between 3 and %d\n",
                                                    command, MAX_T_SIZE);
    exit(-3);
  }

  if( !(t_size % 2) )                            /*ie. t_size is even */
  {
    fprintf(stderr,"%s: template size must an odd number.\n", command);
    exit(-4);
  }

#ifdef OUTPUT_EXTRA_INFO
  fprintf(stderr, ">>Template size: %d\n", t_size);
#endif

/*
 *  If being used, check template coordinates are sensible!
 *  (First adjust coords from centre to top-left corner of template)
 */
  if( !INPUT_TEMPLATE )
  {
    t_xstart -= t_size / 2;  /*Since template-centre coords specified*/
    t_ystart -= t_size / 2;  /*have to adjust these coords to top-left corner*/

    if(t_xstart < 0 || (t_xstart + t_size) > N ||
                                    t_ystart < 0 || (t_ystart + t_size) > N)
    {
      fprintf(stderr,"%s: template coordinates must be between %d and %d\n",
                                    command, t_size/2, N - 1 - t_size/2);
      exit(-5);
    }
```

```
#ifdef OUTPUT_EXTRA_INFO
    fprintf(stderr, ">>Template coordinates: (%d,%d)\n", t_xstart+t_size/2,
                                                         t_ystart+t_size/2);
#endif
  }

/*
 *  Check input and output filenames differ! (only if not stdin/stdout)
 */
  if( infile != NULL && outfile != NULL )
    if( !strcmp(infile, outfile) )
    {
      fprintf(stderr, "%s: input and output filenames identical!\n", command);
      exit(-1);
    }

/*
 *  Check input file access
 */
  if( infile == NULL )                             /*Any infile specified?*/
  {
    f_in = stdin;                                     /*If not use stdin*/
    infile = "stdin";
  }
  else if( (f_in = fopen(infile, "r")) == NULL )   /*Check infile access,etc*/
  {
    fprintf(stderr, "%s: file %s: not found.\n", command, infile);
    exit(1);
  }

/*
 *  Input NxN byte-image into array image[][]
 */
  if(fread(image, N * sizeof(image[0][0]), N, f_in) != N)
  {
    if( feof(f_in) )
      fprintf(stderr,"%s: premature EOF encountered reading from file: %s\n",
                                                             command, infile);
    else if( ferror(f_in) )
      fprintf(stderr, "%s: error reading data from image file: %s\n",
                                                       command, infile);
    exit(-1);
  }

/*
 *  If requested, check template input file access then
 *  input template from it; otherwise extract a section of
 *  the image to act as the correlation template
 */
  if( INPUT_TEMPLATE )
  {
    if( (ftem_in = fopen(templatein, "r")) == NULL ) /*Check access, etc*/
    {
      fprintf(stderr, "%s: template file %s: not found.\n",
                                                  command, templatein);
      exit(1);
    }
    for(y = 0; y < t_size; y++)
    {
      for(x = 0; x < t_size; x++)
      {
        template[y][x] = fgetc(ftem_in);
        if( feof(ftem_in) )
        {
          fprintf(stderr,
```

```
            "%s: EOF encountered reading byte: %d from template file: %s\n",
                                             command, (y*t_size+x) , templatein);
          fprintf(stderr, ">>file contains a template smaller than: %d x %d\n",
                                                                   t_size, t_size);
        }
        if( ferror(ftem_in) )
        {
          fprintf(stderr, "%s: error reading data from template file: %s\n",
                                                            command, templatein);
          exit(-1);
        }
      }
    }
  }
  else
  {
    for(y = t_ystart; y < t_ystart + t_size; y++)
    {
      for(x = t_xstart; x < t_xstart + t_size; x++)
        template[y-t_ystart][x-t_xstart] = image[y][x];
    }
  }

/*
 *  If requested, check template output file access then
 *  output the template to it
 */
  if( OUTPUT_TEMPLATE )                             /*Template outfile specified?*/
  {
    if( (ftem_out = fopen(templateout, "w")) == NULL ) /*Check access, etc*/
    {
      fprintf(stderr, "%s: cannot output template to file: %s.\n",
                                                         command, templateout);
      exit(1);
    }

    for(y = 0; y < t_size; y++)
    {
      for(x = 0; x < t_size; x++)
      {
        if( fputc(template[y][x], ftem_out) == EOF )
        {
          fprintf(stderr, "%s: error writing data to template file: %s\n",
                                                            command, templateout);
          exit(-1);
        }
      }
    }
    fclose(ftem_out);          /* Close the template output file */
  }

/*
 *  Check output file access
 */
  if( outfile != NULL )                               /*Any outfile specified?*/
    if( (f_out = fopen(outfile, "w")) == NULL ) /*Check outfile access,etc*/
    {
      fprintf(stderr, "%s: cannot output to file %s.\n", command, outfile);
      exit(1);
    }

/*
 *  Calculate the mean (dc) value of the template
 */
  templ_dc = 0;
```

```
  for(y = 0; y < t_size; y++)
  {
    for(x = 0; x < t_size; x++)
      templ_dc += template[y][x];
  }

  templ_dc = (int)((double)templ_dc / (double)(t_size * t_size) + 0.5);

/*
 *  Calculate sum-square deviation of the template and also then
 *  subtract the dc term from each element of the template
 */
  templ_ssd = 0;
  for(y = 0; y < t_size; y++)
  {
    for(x = 0; x < t_size; x++)
    {
      templ_ssd += (template[y][x] - templ_dc) *
                                          (template[y][x] - templ_dc);
      template[y][x] -= templ_dc;
    }
  }

/*
 *  Now divide the template through by the root-sum-square-deviation
 *  which is a measure of the template's contrast
 */
  templ_rssd = sqrt((double)templ_ssd);

  for(y = 0; y < t_size; y++)
  {
    for(x = 0; x < t_size; x++)
    {
      norm_templ[y][x] = (double)template[y][x] / templ_rssd;
    }
  }

/*
 *  Go through entire image and correlate template at each x,y position
 *  NB. No allowance is made for edges -- thrown away!
 */
  for(y = t_size / 2; y < N - t_size / 2; y++)
  {
#ifdef OUTPUT_EXTRA_INFO
fprintf(stderr, "%d\r", y);
#endif
    for(x = t_size / 2; x < N - t_size / 2; x++)
    {
/*
 *  Calculate dc term of part of image currently overlaid by template
 */
      local_img_dc = 0;
      for(local_y = 0; local_y < t_size; local_y++)
      {
        for(local_x = 0; local_x < t_size; local_x++)
          local_img_dc += (int)image[y-t_size/2+local_y][x-t_size/2+local_x];
      }

      local_img_dc = (int)((double)local_img_dc /
                                        (double)(t_size * t_size) + 0.5);

/*
 *  Calculate sum-square deviation of image currently overlaid by template
 */
      local_img_ssd = 0;
```

```
      for(local_y = 0; local_y < t_size; local_y++)
      {
        for(local_x = 0; local_x < t_size; local_x++)
          local_img_ssd += (image[y-t_size/2+local_y][x-t_size/2+local_x] -
                                                                 local_img_dc)
                           * (image[y-t_size/2+local_y][x-t_size/2+local_x] -
                                                                 local_img_dc);
      }

      local_img_rssd = sqrt((double)local_img_ssd);

/*
 *  Accumulate sum of the difference squared between the normalised template
 *  and the normalised part of the image; store in corr_val
 */
      corr_val = 0.0;
      for(local_y = 0; local_y < t_size; local_y++)
      {
        for(local_x = 0; local_x < t_size; local_x++)
        {
          imval_dc = image[y-t_size/2+local_y][x-t_size/2+local_x]
                                                              - local_img_dc;
          norm_imval = (double)imval_dc / local_img_rssd;
          diff = norm_templ[local_y][local_x] - norm_imval;
          corr_val += diff * diff;
        }
      }

/*
 *  Place corr_val into current corr_map array element and check whether
 *  this value exceeds the current maximum correlation value: if so update
 */
      corr_map[y][x] = corr_val;
      if(corr_val > max_corr_val)
        max_corr_val = corr_val;
      if(corr_val < min_corr_val)
        min_corr_val = corr_val;
/*
 *  Also update the correlation map overall stats.
 */
      sum_corr_val += corr_val;
      count++;
    }
  }

/*
 *  Evaluate the mean correlation value
 */
  mean_corr_val = sum_corr_val / (double)count;

/*
 *  Use the mean to calculate the s.d. of the correlation map
 */
  for(y = t_size / 2; y < N - t_size / 2; y++)
    for(x = t_size / 2; x < N - t_size / 2; x++)
      sd_corr_map += (corr_map[y][x] - mean_corr_val) *
                          (corr_map[y][x] - mean_corr_val);

  sd_corr_map = sqrt(sd_corr_map / (count - 1.0));

/*
 *  Compute the scale factor to make the correlation map
 *  values go from 0 -> 255
 */
  scale factor = 255.0 / max corr val;
```

```
#ifdef OUTPUT_EXTRA_INFO
  fprintf(stderr, "\nMaximum correlation value = %f\n", max_corr_val);
  fprintf(stderr,"Mean correlation value = %.2f\t", scale_factor*mean_corr_val);
  fprintf(stderr,"  s.d. = %.2f\n", scale_factor*sd_corr_map);
#endif

/*
 *  Now output the thresholded correlation map: use 0 for <= threshold
 *                                                  1 otherwise
 */
  if( outfile != NULL )
  {
/*
 *  Calculate the automatic threshold value for the correlation map
 *  first calculate the parameter which determines the number of s.d's from
 *  the mean at which to place the threshold.  (Note, the three coefficients
 *  have been estimated experimentally).
 */
/*nsds = 2.735 + (0.0733 * t_size) - (0.0014 * t_size * t_size);*Artificial*/
  nsds = 2.6787 + (0.12286 * t_size) - (0.00082 * t_size * t_size);/*MGN subt1*/  th
  if( threshold < min_corr_val )
  {
    fprintf(stderr,"%s: WARNING: correlation map has no regions meeting the match cr
                                            command);
  }
#ifdef OUTPUT_EXTRA_INFO
  fprintf(stderr, "Threshold value chosen: %.1f\n", scale_factor*threshold);
  fprintf(stderr, "(ie. %.4f s.d.'s from the mean)\n", nsds);
#endif

  for(y = 0; y < N; y++)
  {
    for(x = 0; x < N; x++)
    {
      if(y < t_size/2 || y >= N - t_size/2 || x < t_size/2 || x >= N - t_size/2)
      {
        if( fputc(1, f_out) == EOF )
        {
          fprintf(stderr,"%s: error writing to correlation map file: %s\n",
                                                            command, outfile);
          exit(-1);
        }
      }
      else if( corr_map[y][x] <= threshold )
      {
        if( fputc(0, f_out) == EOF )
        {
          fprintf(stderr,"%s: error writing to correlation map file: %s\n",
                                                            command, outfile);
          exit(-1);
        }
      }
      else
      {
        if( fputc(1, f_out) == EOF )
        {
          fprintf(stderr,"%s: error writing to correlation map file: %s\n",
                                                            command, outfile);
          exit(-1);
        }
      }
    }
  }
  }
```

```
/*
 *  Scale correlation values to be a byte-image
 *  First check whether an edge: if so set values to 255
 */
  if( mapfile != NULL )
  {
  unsigned char out_map[N][N];                     /*8bit correlation map */

    if( infile != NULL && !strcmp(infile, mapfile) )
    {
      fprintf(stderr, "%s: input and map filenames identical!\n", command);
      exit(-1);
    }
    if( outfile != NULL && !strcmp(outfile, mapfile) )
    {
      fprintf(stderr, "%s: output filename used twice!\n", command);
      exit(-1);
    }

    if( (f_map = fopen(mapfile, "w")) == NULL ) /*Check mapfile access,etc*/
    {
      fprintf(stderr, "%s: cannot output to file %s.\n", command, mapfile);
      exit(1);
    }
    for(y = 0; y < N; y++)
    {
      for(x = 0; x < N; x++)
      {
        if(y < t_size/2 || y >= N - t_size/2 ||
                                         x < t_size/2 || x >= N - t_size/2)
          out_map[y][x] = EDGE_VAL;
        else
          out_map[y][x] = (unsigned char)(scale_factor*corr_map[y][x]);
      }
    }

    if( fwrite(out_map, N * sizeof(out_map[0][0]), N, f_map) != N)
    {
      fprintf(stderr,"%s: error writing data to map file: %s\n",
                                                           command, mapfile);
      exit(-1);
    }
  }

  exit(0);
}


/*
 *  Function to output `usage:' error message
 */
void
usage(command)
char *command;
{
  fprintf(stderr,
    "usage: %s [-s template_size] [ [-c template_x_coord template_y_coord] | [-tin o
                                                                     command);
  return;
}
```

G.3.2 C program file "SIMPLE_CORREL" - Dr. D. Greenwood

(N.B. this version of the program does not include normalisation).

```
SIMPLE_CORREL.c      Mon Jul 15 10:16:22  1991

/*
 *      compute correlation of two image regions and output resulting
 *      correlation surface. Movement of the mask window within the search region
 *      can be subpixel, and the type of resampling (transformation) of the
 *      right (mask image is simply x/y translation. Search window values are
 *      output for plotting.
 *
 *      d.greenwood         jul.91
*/

#include <stdio.h>
#include <match.h>
#include <pixrect.pixrect_hs.h>

#define REAL    double              /* float or double ?? */
#define ABS(_t)             ((_t) < 0 ? -(_t) : (_t))

#define result_file stdout
#define plot_file   stdout

Pixrect *left, *right;  /* image file pixrects */
int     lx, ly, rx, ry; /* starting positions (origins) */
REAL    Search_Radius;  /* size of search area */
int     Match_Radius;   /* size of mask/correlation window */
REAL    Step;           /* incremental movement of mask window */

REAL    **Correl;       /* holds correlation values */
REAL **New;             /* holds resampled right image */

int     Mode;           /* apply different parameter sets in resample' */

REAL    Delta_param1;   /* param1 at minimum */
REAL    Delta_param2;   /* param2 at minimum */

/*
** main() - the main program,
*/
main(argc, argv)
  int argc;
  char  **argv;
{
  int i, j;
  int msize, size;

        parse(argc, argv);

        /*
        ** workspace memory
        */
        size = (int) (Search_Radius * 2.0 / step + 0.5) + 1;
        msize = (Match_Radius * 2) + 1;

        Correl = (REAL **)calloc((unsigned)size, sizeof(REAL *));
```

```
        for(i=0; i < size; i++)
                Correl[i] = (REAL *)calloc((unsigned)size, sizeof(REAL));

        New = (REAL **)calloc((unsigned)msize, sizeof(REAL *));
        for(i=0; i < msize ; i++)
                New[i] = (REAL *)calloc((unsigned)msize, sizeof(REAL));

        if(correl == NULL || New == NULL) {
                fprintf(stderr,"memory allocation failed - exit\n");
                exit(-1);
        }
        fprintf(stderr,"working");

        /*
        ** compute correlation
        */
        correlate(left, right, lx, ly, Search_Radius, Match_Radius, Step);

        fprintf (stderr,"\ndone. \n");
        fprintf (stderr,"\nminimum @ <%f> <5f>\n", Delta_param1, Delta_param2);
        /*
        ** output correlation surface
        */
        fprintf (stderr," %d plot values\n, size, size);
        for (j=0; j< size; j++)  {
        for(i=0; i < size; i++){
        fprintf(plot_file,"%f ", Correl[j] [i]);
        }
        fprintf(plot_file,"\n");
        }
        free(Correl);
        free(New);
        exit(0);
}
/*
** interpolate:
**
**      bilinear interpolation of the images at (x,y)
*/
REAL interpolate(image, x, y)
  Pixrect *image;
  Real x, y;
{
  register int ix, iy, offset;
  register REAL fx;
  register REAL t1, t2;
  REAL  ret;
  register unsigned char *ptr;

        offset = (mpr_d(images)->md_linebytes*8/images->pr_depth)-images->pr_width;
        ix = (int)x;
        iy = (int)y;
        fx = x - ix;
        ptr = (unsigned char *)mpr_d(image)->md_image) + iy*(image->pr_width) = iy*offset + ix;
        t1 = (REAL) (*ptr) + fx*((REAL) (*(ptr+1)) -REAL) (*ptr)):
        ptr += (image->pr_width + offset);
        t2 = (REAL) (*ptr) + fx*((REAL) (*(ptr+1)) -REAL) (*ptr)):
        ret =t1 + (y - (REAL) (iy))*(t2 - t1);
        return(ret);
}
```

```
/8** resample() - resample the right image for the given 'Mode'.
**
*/
resample(image, origin x, origin y, param1, param2, radius)
  int origin_x, origin_y;
  REAL param1, param2;
  register int radius;
{
  register int i,  j;
  register REAL xpos, ypos;

        for(j=(-radius); j <= radius; j++)  {
                for(i=(-radius);  i <= radius; i++)  {
                        switch(Mode)  {
                        case 1:
                                /* X-trans vs Y-trans */
                                xpos = (REAL)origin_x + (REAL)i + param1;
                                ypos = (REAL)origin_y + (REAL)j + param2;
                                break;
                        case 2:
                                /* X-trans vs Y-shear */
                                xpos = (REAL)origin_x + (REAL)i + param1 + ((REAL)j)
                                ypos = (REAL)origin_y + (REAL)j;
                                break;
                        case 3:
                                /* Y-trans vs X-shear */
                                xpos = (REAL)origin_x + (REAL)i + ((REAL)j) * param2);
                                ypos = (REAL)origin_y + (REAL)j + param1;
                                break;
                        case 4:
                                /* X-trans vs Y-shear */
                                xpos = (REAL)origin_x + (REAL)i + param1;
                                ypos = (REAL)origin_y + (REAL)j + (REAL)i * param2);
                                break;
                        case 5:
                                /* Y-trans vs Y-shear */
                                xpos = (REAL)origin_x + (REAL)i;
                                ypos = (REAL)origin_y + (REAL)j + param1 + ((REAL)i * param2);
                                break;
                        case 6:
                                /* X-shear vs Y-shear */
                                xpos = (REAL)origin_x + (REAL)i + (REAL)j * param1);
                                ypos = (REAL)origin_y + (REAL)j + (REAL)i * param2);
                                break;
                        default:
                                xpos = (REAL)origin_x + (REAL)i;
                                ypos = (REAL)origin_y + (REAL)j;
                        }
                        if((xpos < 0.0)  ||  (xpos >=  (REAL)image->pr_width)
                        || (ypos < 0.0)  ||  (ypos >=  (REAL)image->pr_height))  {
                                fprintf(stderr,"error - resample out of range of rig")
                                fprintf(stderr,"origin_x:%d  origin_y:%d param1:%f")
                                        origin_x, origin_y, param1, param2, i, j);
                                exit(-1);
                        }
                        New[j+radius][i+radius] = (REAL)interposlate(image, xpos, ypos);
                }
        }
/*
fprintf(stderr,"\nresample: shift(%f, %f)\n", shift_x, shift_y);
print_array(New, (2*radius + 1),  (2*radius + 1));
```

```
*/
}
** correlate() -          for each point in the search window (centered at (rx,ry)),
**                        compute the least square correlation between the template
**                        and the mask window and store in the result array.
*/
correlate(left, right, lx, ly, rx, ry, search_rad, match_rad, step)
  Pixrect *left, *right;
  int lx, ly, rx, ry;
  REAL  search_rad;
  int   match_rad;
  register REAL step;
{
  register REAL           Sum = 0;
  register REAL           Stemp;
  register REAL           i,  j;

  register int            k,  l,;
  register int            x,  y;

  REAL  minimum=1e100;
  int   timer=0;

        /*
        ** mask window moved over search window in 'step' increments
        ** for the given type of resampling - eg. translation vs shearing.
        ** 'step' may be fractional, hense i and j are real.
        */
        for (j = (0.0-search_rad); j < (search_rad+step); j+step)  {
                fprintf(stderr,"\n\t");
                   for (i = (0.0-search_rad); i < (search_rad+step); i+step)  {
                        fprintf(stderr,"."); fflush(stderr);

                                 /*
                                ** resample right image using a combination
                                ** of image distortions - eg. i= x translation,
                                ** j= y shearing.
                                */
                                if(i < ).)-search_rad)  ||  i > (search_rad+step)  || j < (0.0-
search_rad)

                                        fprintf(stderr,"i and j corrupted before call to resample");
                                resample(right, rx, ry, i, j, match_rad);

                                Sum = 0.0;

                                /*
                                ** 'New' has been resampled from right image, so now
                                ** just do least-squares summation over mask window.
                                */
                                for (k = (-match_rad); k <= match_rad; k++)  {
                                        for (l = (-match_rad); l <= match_rad; l++)  {
                                                temp = ((REAL)pr_get(left, lx+l, ly+k)) - New;

                                                Sum +=  (REAL) (temp * temp);
                                        }
                                }
                                Sum /= (REAL) (2*match_rad = 1)*(2*match_rad + 1));


                                /*
                                ** convert to integer indices and store Sum.
```

```
                                */
                                x = (int) (i+search_rad)/step + 0.5);
                                y = (int) (j+search_rad)/step + 0.5);
                                if (Sum < minimum)  {
                                        minimum = Sum;
                                        Delta_param1 = i;
                                        Delta_param2 = j;
                                }

                   }
         }
}

/*
** parse() - very basic command line parse + parameter grabbing.
**                Limited error checking!
*/
parse(argc, argv)
  int argc;
  char **argv;
{
  FILE  *infile,  *fopen();
  Pixrect  *pr_load();
  float tempf;
  int temp,  size;
  unsigned char  *ptr;

        if (argc  != 3)  {
                fprintf(stderr,"format: %s left_image right_image\n", argv[0]);
                exit(1);
        }
        /* load left image */
        if((infile=fopen(argv[1], "r")) == NULL)  {
                fprintf(stderr,"unable to open %s\n", argv[1]);
                exit(1);
        }
        if((left=pr_load(infile, NULL)) == NULL)  {
                fprintf(stderr,"unable to load %s\n", argv[1]);
                exit(1);
        }
        fclose(infile);

        /* load right image */
        if((infile=fopen(argv[2], "r")) == NULL)  {
                fprintf(stderr,"unable to open %s\n", argv[2]);
                exit(1);
        }
        if((right=pr_load(infile, NULL)) == NULL)  {
                fprintf(stderr,"unable to load %s\n", argv[2]);
                exit(1);
        }
        fclose(infile);

        fprintf(stderr, "number of significant bit-planes:  ");
        fflush(stderr);
        scanf("%d, &temp);

        temp = 8 - temp;
        if(temp > 1)  {
                size = left->pr_width * left->pr_height;
                ptr = (unsigned char *)mpr_d(left)->md_image;
```

```
                rop_fastloop(size,  (*ptr)  = (*ptr)>>temp;  ptr++);

                size = right->pr_width * right->pr_height;
                ptr = (unsigned char *)mpr_d(right)->md_image;
                rop_fastloop(size,  (*ptr)  = (*ptr)>>temp;  ptr++);

                fprintf(stderr,"left  (template)  image  (x,y):  ");
                fflush(stderr);
                scanf("%d  %d",  &lx,  &ly);                      /* match point */
                fprintf(stderr,"right  (mask)  image  (x,y):  ");
                fflush(stderr);
                scanf("%d  %d",  &rx,  &ry);                      /* initial seed */

                fprintf(stderr,"mask window radius :  ");
                fflush(stderr);
                scanf("%d",  &Match_Radius);

                fprintf(stderr,"parameter range  (pixels) : ");
                fflush(stderr);
                scanf("%f",  &tempf);
                Search_Radius = (REAL)tempf;

                fprintf(stderr,"parameter sub-pixel increment (0.0 <-> 1.0) : ");
                fflush(stderr);
                scanf("%f",  &tempf);
                Step = (REAL)tempf;

                fprintf(stderr,"Parameter set:\n");
                fprintf(stderr,"\t(1) : X-translation vs Y-translation\n");
                fprintf(stderr,"\t(2) : X-translation vs X-shearing\n");
                fprintf(stderr,"\t(3) : Y-translation vs X-shearing\n");
                fprintf(stderr,"\t(4) : X-translation vs Y-shearing\n");
                fprintf(stderr,"\t(5) : Y-translation vs Y-shearing\n");
                fprintf(stderr,"\t(6) :    X-shearing vs Y-shearing\n");
                do  {
                                fprintf(stderr,"Choose: ");
                                scanf("%d", &Mode);
                                if(Mode < 1  || Mode > 6)
                                        fprintf(stderr,"\t  *invalid selection*\n);
                } while  (Mode < 1  || Mode > 6);

                fflush(stderr);

                temp = Match_Radius + ((int) Search_Radius);

                if(((lx-temp) < 0)  || ((ly-temp) <0)
                || ((lx+temp) > (left->pr_width-1))  ||  ((ly+temp)  >  (left->pr_height-1)))  {
                        fprintf(stderr,"error - mask/search windows will wander off image  -
                        exit(-1);
                }
                return;
}

/*
** print_array(array() - debug dump of REAL array.
**
*?
print_array(array, xsize, ysize)
  REAL **array;
  int xsize, ysize;
{
```

```
  register int i, j;

                fprintf(stderr, "&d %d\n", xsize, ysize);
                for(j=0; j < ysize; i++)  {
                        for(i=0; i < xsize; i++)
                                fprintf(stderr,"%f ", array[j][i]);
                        fprintf(stderr,"\n");
                }
}

/*
** tic_tic()
**      make cursor look like rotating clock hands
*/
tic_tic(pos)
  register int *pos;
{
                switch(*pos)  {

                case 0:         fprintf(stderr, "¦\b");
                                fflush(stderr);
                                *pos = 1;
                                break;
                case 1:         fprintf(stderr, "¦\b");
                                fflush(stderr);
                                *pos = 2;
                                break;
                case 2:         fprintf(stderr, "¦\b");
                                fflush(stderr);
                                *pos = 3;
                                break;
                case 3:         fprintf(stderr, "¦\b");
                                fflush(stderr);
                                *pos = 0;
                                break;
                }
}
```