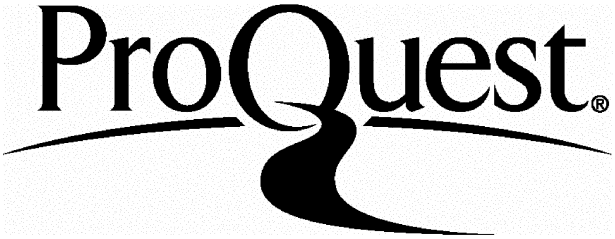
Registration and quantitative comparison of temporal mammograms (with application to HRT data)

Konstantinos Marias
BSc, MSc

A thesis submitted in partial fulfilment of
the requirements for the PhD degree

2002

Registered at
Royal Free Campus and University College Medical School
University College London



Abstract

In order to reduce the high mortality rate of breast cancer, most research in mammogram image analysis aims to develop a CAD (Computer Assisted Detection) system that can assist the clinician in the difficult task of "early" diagnosis. The comparison of temporal pairs of mammograms is believed by radiologists to be often crucial for diagnosing cancer, especially as breast tissue is highly variable across the population making it difficult to perform diagnosis reliably from a single mammogram. However, changes in the appearance of mammograms due to differences in compression and imaging conditions can limit the effectiveness of temporal comparison. In this thesis, a registration method that aligns temporal (or bilateral) mammograms is presented. This technique has the potential to assist the clinician to detect changes more efficiently (e.g. interval cancers in mammogram sequences, architectural distortions or microcalcifications in bilateral mammograms). In addition, the application of this technique to mammograms of Hormone Replacement Therapy (HRT) users is investigated. Since long-term use of HRT can increase the risk of breast cancer (as a side effect of glandular tissue regeneration), quantification of temporal tissue density changes (in addition to registration) is needed for assessing density changes *locally*. In order to derive quantitative measures of breast tissue change, the mammogram pairs are processed using the h_{int} representation of interesting tissue. Since such a measure depends on the image information context, we examine the problems that arise from combining image registration and quantification, aiming to develop a robust framework for temporal tissue density change assessment.

Acknowledgements

I would like to thank the Mike Brady for being always there when I needed him, either as a supervisor or as a friend. His words never failed to put a smile in my face and to give me courage.

I would also like to express my sincere gratitude to my supervisors Alex Seifalian and Santilal Parbhoo as well as Gloria Freilich from Cancerkin Centre, Royal Free Hospital, for supporting my PhD. Special thanks to Alex for his constant moral support and guidance and to Mr. Parbhoo for allowing me to attend his clinics during my first year and helping me understand the clinical problems related to mammography.

I am indebted to Chris Behrenbruch and Ralph Highnam. Chris has been a great friend and a great colleague. I really enjoyed working and socialising with him. Ralph gave me lots of valuable and robust advice all these years.

Many people from the Medical Vision Laboratory in Oxford have been great colleagues and friends all these years. To Gerardo, Miguel-Angel, Marius, Miguel, Seb, Patrick, Maud, Albert, Rob, Paul, Michael, Margaret, Guo-fang and Carolyn thank you for your support and friendship. In particular, thanks to Paul and Carolyn for reviewing my thesis.

Outside work, lots of people stood by me and helped me in difficult times. To Jim, Susana, Enrique, Anne, Doros, Christina, Anastasia, Eleana, Adriani and Georgiana thank you for caring for me.

Finally, I would like to thank my beloved family. My mother Ioanna, my father Teri, my brother Vasili, my sister Maria, and my grandfather Vasili, to whom I promised a long time ago, I would never forget him. And I didn't.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Introduction to Medical Image Analysis

Over the past two decades, medical image analysis has employed several image processing and computer vision concepts to enhance, describe and process digital images in order to improve their clinical utility. This is a part of a larger scale application of technology to medicine that has been inspired by the development of increasingly powerful computers and has made a huge impact on modern clinical medicine.

Medical images of human organs can assist a clinician in diagnosis, surgical planning and intervention and evaluation of therapy. Furthermore, functional imaging as well as histological imaging, can assist the exploration of complex physiological (or pathological) processes in the human body. Some specific areas of research include:

- Surgical planning and guidance: Preoperative data are used for planning the surgery more efficiently; the surgeon performs the operation guided by image data acquired at the time of the operation.

- Evaluation of therapy: This involves the accurate comparison of temporal data for evaluating the effectiveness of a specific therapy against a disease.

- Tele-medicine applications:
  - Remote diagnosis: The expert can establish diagnosis (or give a second opinion) by examining medical data from remote locations.
  - Tele-surgery: surgery that is supervised or aided by a surgeon located remote from the scene of the operation.

- Assisted diagnosis through:
  - Enhancement and characterisation, to remove noise or to manipulate the contrast characteristics of a region of interest and to describe the texture (e.g. homogenous vs. high entropy) of that region.

- Comparison of temporal data of the patient, or data from different modalities.
- Anatomical or physiological "models" of normal variation, in order to automatically detect regions that potentially exhibit an abnormality.
- Clinical databases for differential diagnosis.

The specific motivation for this thesis is breast cancer. The main objectives of breast cancer imaging are:

- Early diagnosis of the disease and reduction of the number of missed cancers (false negatives) so that mortality rates reduce.
- Reduction of the number of false positives that often lead to unnecessary biopsies and consequently to psychological trauma for the patient.
- Analysis of the image changes that correspond to the effects of therapy for assessing the success of a specific drug in treating the cancer.

A detailed account of breast anatomy, physiology and pathology can be found in Appendix A of the thesis.

Medical image analysis has provided sophisticated imaging techniques which may assist the interpretation of radiological images. It has a crucial role to play in providing the clinician with objective anatomical and quantitative information directly from mammograms. In this chapter, the role of the breast-screening programme is discussed with special emphasis in the need for image analysis tools to improve its effectiveness, and the key components of the thesis are summarised.

1.2 Breast cancer and the screening programme

1.2.1 Introduction

In the western world, breast cancer is the most common cancer among women. In the European Community (EC) breast cancer represents 19% of cancer deaths and 24% of all

cancer cases. It is the primary diagnosis in a total of 157,000 cases annually and kills almost 70,000 annually [1]. As already mentioned in the introduction of this chapter, technological advances in breast cancer imaging are aimed at early diagnosis. Although mammography remains the predominant modality for breast cancer screening and diagnosis, an increasing number of modalities (e.g. contrast-enhanced Magnetic Resonance Imaging (MRI), Scintimammography) have been used in clinical practice (as is discussed in the next chapter). However, technological advances in breast imaging alone cannot tackle the problem; government-defined, global health strategies for the early diagnosis of the disease are required as well. Over the last 40 years, the ideas of using the technology for the early detection of breast cancer in a whole population rather than on an individual-basis (and very often post-symptomatic) assessment, have evolved, particularly in so-called screening programmes.

The first large-scale trial undertaken in New York (between 1961 and 1967) was successful as women who had a cancer detected had a much higher probability of the lesion being pre-invasive [2] and hence good prognosis. However, it was not until 1976, when the first screening program with randomised, controlled, population-based trials was introduced in Sweden [3]. It demonstrated a nearly 30% reduction in mortality against women that did not participate and many European countries followed this example.

In The United Kingdom, the design of the first screening programme was undertaken by a working group under Sir Patrick Forrest, whose report was accepted by the government in 1986. X-ray mammography was proven to be a cost effective imaging modality for national screening. However, mammography is not diagnostically useful for younger women as their breasts consist mostly of glandular tissue which often masks masses (both benign and diseased tissue have very "bright" mammographic appearance). As a consequence, the UK screening programme was established for women between 50 and 64. The implementation of the screening programme has involved the establishment of specialist breast screening centres and formal training of both radiographers and radiologists.

This led to several implications for hospital radiology departments. Before the implementation of the screening programme, the detection of masses was as a result of women reporting to their general practitioner with a breast lump, breast pain or other pathological indicators. In most cases these women were referred directly to the symptomatic unit where the mammograms were examined by radiologists who were not always very experienced in reading mammograms. Today, in the UK, many specialist screening centres are staffed by radiologists specialised, often exclusively, in screening mammography. However, there is a critical shortage of radiologists trained to read mammograms while the amount of patient data is steadily increasing.

Because of the screening programme, small and usually non-palpable masses are increasingly detected. As a result, the radiologist's task has become more complex since the interpretation of the mammogram has becomes more challenging. A cancer that appears in the screening interval is known as an interval cancer. The large number of interval cancers leads to the conclusion that the Forrest report was over-optimistic [4].

1.2.2 A short analysis of the Screening Programme.

Breast screening programmes were introduced in order to reduce the overall mortality rate. Screening is defined as the process of inviting a specific population to undergo particular examinations which may detect whether or not a given disease is present [2,3]. In general, a breast-screening programme is based upon combinations of mammography and clinical examinations. The main parameters that can define the effectiveness of the screening programme are:

- **Design of the screening programme**. This includes defining the age group to be invited, design of the clinical process (e.g. first mammogram acquisition then reading followed by recall if necessary) and the necessary infrastructure and cost-analysis.

- **X-ray image-acquisition and quality**. This is usually a trade-off with patient radiation dose.

- **Proper-positioning and breast compression**. The entire breast should be imaged (including the axilla) and compression should be optimal in terms of glandular-tissue separation.

- **Level of medical expertise**. The early detection of a lesion is *strongly* dependent on the experience of the radiologist that reads the mammogram.

- **Consistency**. For a screening programme to be successful, it is necessary to ensure that the clinical protocol is consistent in order to facilitate individual and population comparison. This is very challenging to ensure in practice.

In the typical scenario, mammograms are acquired in regional screening units and subsequently examined in reading centres where the decision to recall a woman for further assessment is made. In the UK, more than 200 radiologists are employed in the national screening programme and the expected screening output is around 5000 cases a year [5]. When reading a patient's mammograms, the clinician compares them with previous ones in order to detect changes. This fact is an important motivation for this research as is discussed in the next chapter. Temporal comparison is an important aspect of future CAD (Computer-aided detection or diagnosis) systems that are expected to improve the cancer detection rate.

In summary, breast screening is a cost effective technique that has contributed to the early detection of breast cancer as small, non-palpable masses can be detected sooner. However, greater clinical efficiency and skill is required to deal with the large amount of data particularly as a significant number of cancers are missed. The rapid development of digital mammography has further encouraged the development of computer algorithms that can provide objective information about individual mammograms. To date, most of the researchers in mammogram image analysis (including the research described in this

thesis) still use *digitised* mammogram images (films are usually digitised using a laser scanner).

It is hoped that future use of software will improve the effectiveness of the screening programme, mainly by assisting less experienced radiologists in the early detection of breast cancer and by reducing the variability in standards of the in-service performance. **The work presented in this thesis is concerned with the development of software tools that allow a more efficient comparison of temporal or bilateral mammographic data. More specifically, the aim is to develop a technique to align mammogram sequences.** The motivation for this work is detailed in the next sectioned while section 1.4 summarises the work on temporal mammogram registration. As will be discussed later in section 1.5, this method can be adapted for local quantitative analysis of HRT (Hormone Replacement Therapy) temporal mammogram pairs.

1.3 Motivation and aims of our work

Most of this work is motivated by the reported limitations of the screening programme in cancer detection. It has been reported that 27% of the missed interval cancers (cancers that "appear" within the screening interval) could have been detected in the previous mammogram [6] and that many of such missed diagnoses are attributed to observer errors [7]. Our work in mammogram registration aims to assist the clinician to detect differences in the comparison of temporal data when reading the mammogram. The results of the thesis are currently being exploited under licence from Oxford University by a software company (http://mirada-solutions.com).

As mentioned previously, there is a need for reliable medical image analysis tools that can assist radiologists (especially the less experienced) to read mammograms and detect subtle changes. In each screening visit, the radiologist usually compares the current mammogram of a patient with previous ones in order to detect changes that could make explicit the presence of an abnormality. Since the number of women returning for a

second, and third, round of breast screening is increasing rapidly the need to compare mammograms becomes more pressing.

The work on temporal mammography aims to provide a computer-based framework for a more effective and intuitive comparison of temporal (or bilateral) data. Section 1.4, further explains the nature of the work on *temporal mammography.*

The second clinical problem that motivated the work presented in this thesis, is the controversy over the increased risk of breast cancer for Hormone Replacement Therapy (HRT) users. It has been suggested that hormone-stimulated tissue regeneration can in some cases lead to the development of an abnormality. Changes in breast-tissue density are crucial for assessing the response of the woman to therapy, while knowledge of the distribution of such changes can assist the clinician to determine whether there is an increased risk of cancer for the *individual woman.* To date, it is not possible to obtain quantitative information from mammograms. The objective is to develop tools that can facilitate the quantitative comparison of HRT mammogram sequences. HRT data are considered as a special application of temporal mammography, where there is an additional need to *quantitatively* compare each mammogram pair. In section 1.5 a short account of the work on HRT data is provided.

1.4 Temporal mammography

Mammogram images are highly variable. For this reason it is currently impossible to define a consistent method for comparison. Worse, clinical signs are often subtle. For these reasons, the radiologist's most powerful weapon is *change* detection. As was discussed in the previous section, screening programs have been implemented in many European countries, (the UK, Sweden, Finland and the Netherlands being the major examples) as an attempt to reduce mortality attributed to breast cancer through early diagnosis. This has also affected clinical diagnosis since temporal sequences of mammograms started to be available for individual women, a fact that gave clinicians an additional advantage for breast cancer diagnosis: temporal comparison of x-ray

mammograms of the same patient. Since screening re-visits are increasing, more temporal data is becoming available to the clinician, making temporal comparison of mammograms an absolute necessity in clinical practice. Similar to bilateral comparison, the clinician compares the most recent mammogram with previous ones in order to detect abnormalities on the basis of "significant" differences. In the remainder of this thesis this process will be called *temporal mammography*. As was mentioned before, the main objective of this work is the development of tools for the alignment of temporal mammogram.

In Figure 1, temporal sequences of the same patient spanning 8 years are shown. Such sequences provide considerable information about the temporal changes in breast tissue and can assist diagnosis, since any significant difference in the architecture of the breast tissue over time can be considered as a candidate region for a malignancy. In Figures 1 (d) and (e), the region pointed with arrows could have been confused with a recurrent cancer. In Figure 2, however, after registering all the previous images (with the method that is described in chapter 4), it becomes obvious that these regions correspond to scar tissue from previous surgery (as is highlighted by the arrows).

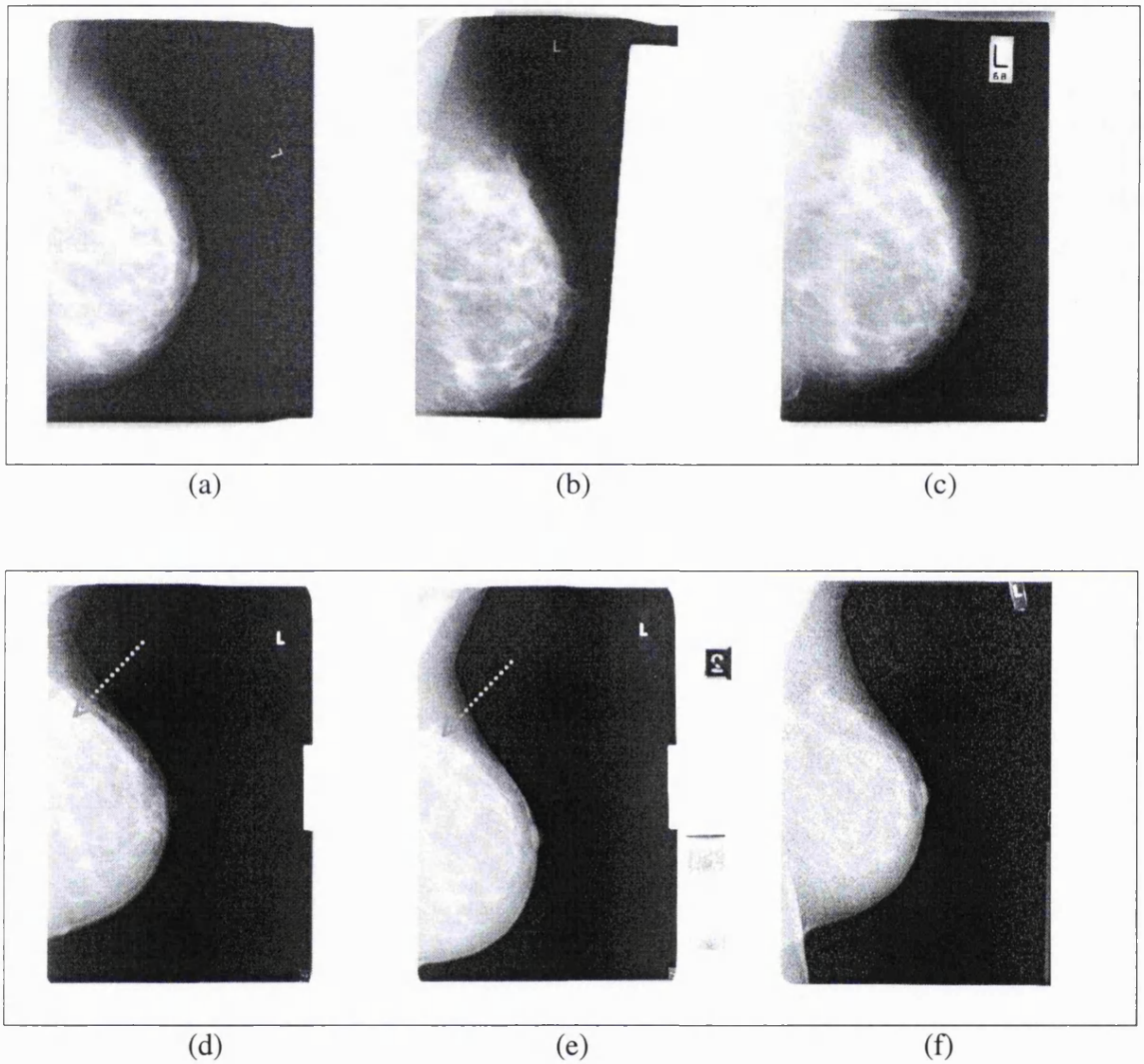


Figure 1: Series of temporal mammograms of the same patient, over a period of 8 years: (a) the oldest and (f) is the most recent. Breast cancer was diagnosed in (c) in the Upper Outer Quadrant (UOQ) of the breast, and was excised. Looking at the images from (a) to (f), it can be noticed that the architecture of the breast-tissue is preserved although the tissue is displaced non-rigidly between images. The reduction of brightness in several areas due to involution is also noticeable. In (d) and (e) there is a bright region (see arrows) that appears as a cancer.

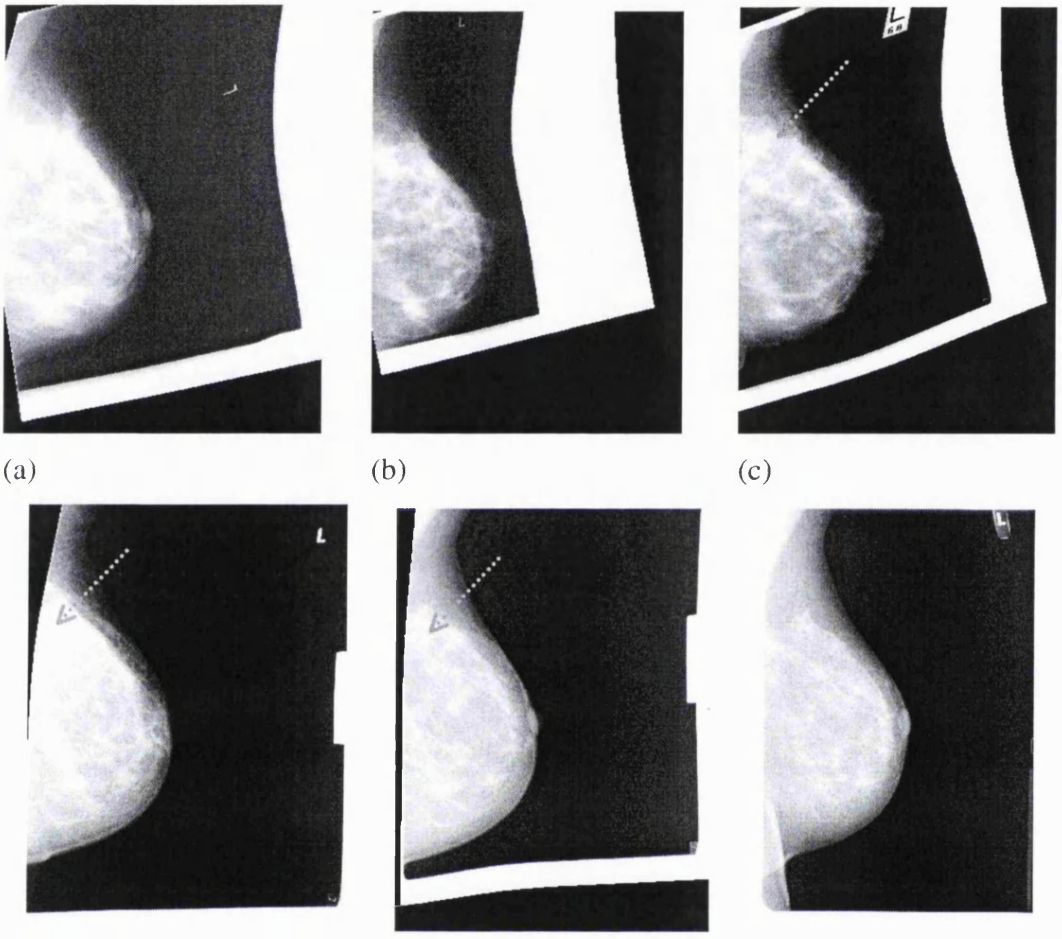


Figure 2: The same series of temporal mammograms (as in Figure 1) registered to the same co-ordinate frame of the most recent. (a) the oldest and (f) is the most recent one. Breast cancer was diagnosed in (c) in the Upper Outer Quadrant (UOQ) of the breast, and was excised. Looking at the images from (a) to (f), it can be seen that the architectural similarity is more pronounced when looking at the registered mammograms. This way, it becomes easier to follow the evolution of regions over time and it becomes easier to understand that the bright region in (d) and (e) is scar tissue after the cancer in the upper outer quadrant in (c) was excised (see arrows).

1.5 Quantitative comparison of HRT mammograms

A topical and increasingly important clinical application of temporal mammography is the analysis of mammogram sequences of Hormone Replacement Therapy users (HRT). As is discussed in detail in chapter 5, HRT can stimulate glandular tissue regeneration, which results in an increased tissue density. As a consequence, it is possible to develop breast cancer while diagnosis becomes more difficult due to the overall increase in brightness.

The first objective was to apply the temporal registration techniques developed to monitor changes wrought by HRT. Working closely with experienced clinicians, several issues that are of specific importance in temporal HRT data were identified:

- If the HRT user responds to therapy, the tissue density and/or the mammographic pattern can change significantly. In this case, as discussed in chapter 5, the HRT-induced pattern should be used as the new baseline for temporal comparison.

- Since cancers in HRT users are related to the hormone-stimulated increase in fibroglandular tissue, it is crucial to be able to monitor such changes in density. This becomes even more necessary since mammographic sensitivity can decrease due to the increased density.

- Changes due to HRT can be characterised either on a global or on a local basis. Global changes represent the overall response to exogenous hormones, while local changes reflect the fact that breast-tissue receptivity to hormones is variable. From a more clinical (rather than theoretical) point of view, certain locations in the breast have a higher frequency of cancer appearance (e.g. the upper outer quadrant) and it is believed that they exhibit a higher activity and are therefore more likely to respond to hormones.

From these observations, it became apparent that image analysis tools could influence clinical decision making for HRT users and could assist the clinician to assess the risk of cancer for *an individual* woman using HRT. Since it is recommended that the first HRT mammogram be used as a baseline mammogram, new growths may be detected on the basis of significant differences with previous mammograms (via registration). Figure 3, shows a temporal HRT mammogram pair. After registration (using the method that is described in chapter 4), two new regions are pronounced in the difference image (Figure 3 (d)). The first corresponds to a density increase while the second corresponds to a cancer.

HRT mammogram sequences may be considered as a special case of temporal mammography where the increased risk of cancer requires a careful study of the changes in breast tissue density. For this reason, further to the work on temporal registration, a temporal quantification scheme was necessary to enable the clinician to assess density changes in a woman that responds to HRT. The quantitative measures proposed were based on the h_{int} representation of interesting tissue, developed by Highnam and Brady [4] (summarised in Appendix C of the thesis). This way, images are normalised for differences in imaging conditions and each pixel value of the h_{int} image can be correlated to the underlying anatomy. To validate the proposed measures of density change, they are compared to the clinician's assessment of tissue changes in 60 HRT temporal pairs.

In chapter 5, the potential to use registration for quantification of local changes is also investigated. Combination of temporal registration with the h_{int} representation could enable the quantitative assessment of local density changes. To overcome the change in image "volume" (i.e. the sum of the h_{int} values in the image) due to registration, a method that preserves the "volume" by calculating an intensity-correction field directly from the calculated transformation was developed.

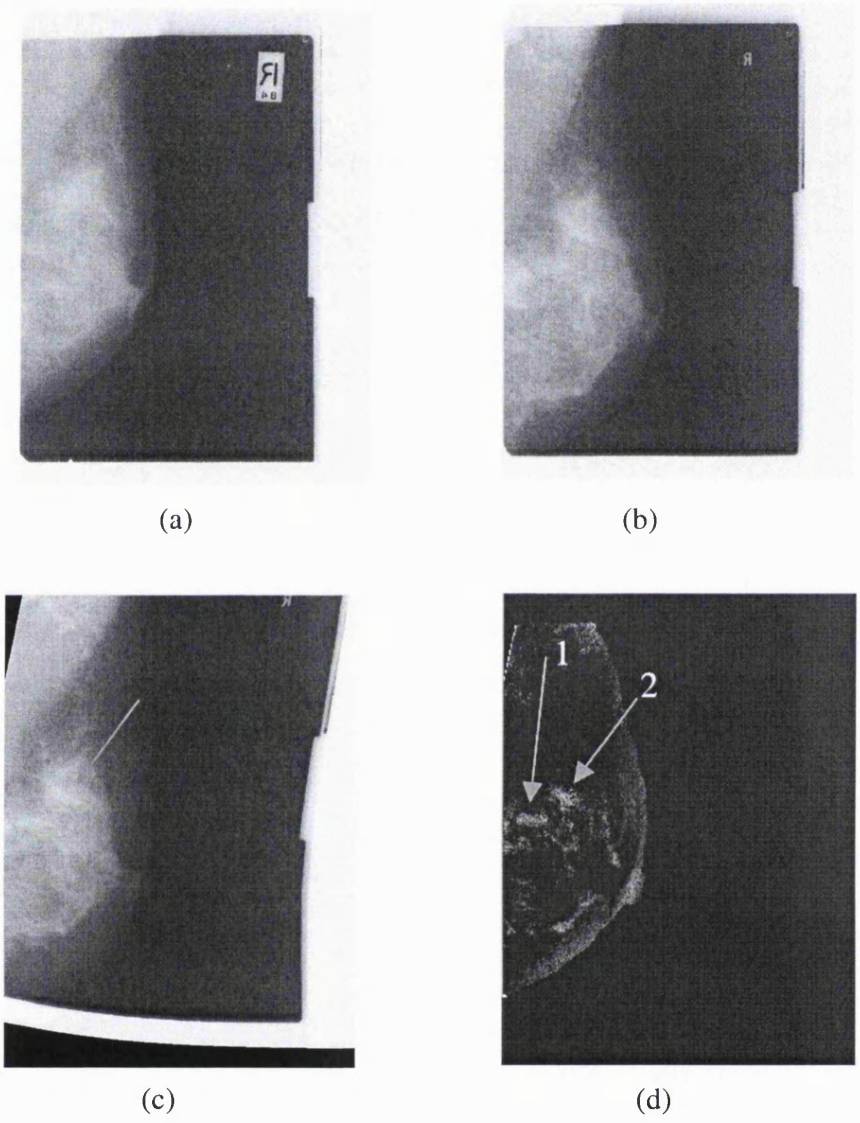


Figure 3: (a), (b): An HRT mammogram pair. In (c), the recent mammogram (b), is registered to the previous one (a). The difference image in (d) highlights 2 new regions. The first one corresponds to a density increase (arrow "1" in (d)), while the second one to a cancer (arrow "2" in (d))

1.6 Overview of the thesis

Chapter 2 summarises the main applications of computer vision in mammography, highlighting the significant role that CAD (computer aided diagnosis) tools may play in the future. After describing the limitations of X-ray mammography, the main characteristics of the most important imaging modalities that are used today in clinical practice are presented, stressing the need to combine information not only from the same modality over time (e.g. temporal mammogram registration) but also from different modalities (data fusion).

Chapter 3 is an analytical review of the registration problem in medical image analysis. The different categories of registration applications are discussed, and some of the more popular techniques that have been developed to implement registration in medical image analysis are summarised. After describing the specific characteristics and requirements of mammogram registration, the relevant previous work on mammogram matching and registration is presented.

Chapter 4 is a description of the proposed method for mammogram registration. The robust estimation of boundary landmarks is explained and mammograms are first aligned using only the boundary. Providing that the breast is not fully involuted*, a multi-scale technique to segment regions of dense tissue inside the breast is used. Matching these regions and including their centroids as internal landmarks can improve registration. Validation results are presented as well as results on registration showing the possible applications of the technique.

Chapter 5 is a description of the work on HRT sequences. First, a detailed review of clinical studies is presented, in order to understand the importance of the clinical problem as well as the controversy concerning several issues related to the use of HRT. Then the potential of the proposed work in influencing clinical decision-making and assisting the clinician in assessing the risk of cancer for the individual woman using HRT, is

* *Involution* is a physiological process that describes the gradual degradation of the dense breast tissue to fat

discussed. Validation results for the proposed measures of global tissue-density change are presented. Finally, a novel method that can potentially enable the combination of image registration without significantly changing the image volume (sum of the h_{int} values) is proposed and a validation experiment is presented.

Chapter 6 summarises the work presented in the thesis and makes explicit its limitations. Then, several suggestions for improving the methods proposed are discussed. Finally, some preliminary results on texture segmentation are presented and the thesis concludes with a discussion on some recent advances in breast imaging data fusion between MRI and x-ray mammography.

Appendix A is an overview on breast anatomy, physiology and pathology

Appendix B is a description of the multi-scale segmentation technique that was applied in chapter 4 to segment internal regions in each mammogram.

Appendix C is an overview of the h_{int} representation that is used in chapter 5 to normalise a mammogram pair and to build reliable quantitative measures of tissue-change.

Appendix D is a summary of the most common definition of terms related to diagnostic tests. The aim of this appendix is to help the reader better understand several references to clinical studies included in the thesis.

Appendix E is a gallery of results of mammogram registration. These include temporal mammograms, interval cancers, bilateral mammograms and HRT pairs. For each pair the transformed image is shown, as well as the subtraction image after registration and in certain cases the joint histograms and additional notes to describe the importance of the registration result.

Chapter 2: Medical image analysis in mammography

2.1 Introduction

In this chapter, the various applications of medical imaging to breast cancer are discussed, emphasising the specific areas where our work aims to contribute. The main role of medical images is their use in patient management, starting from diagnosis (e.g. breast screening using X-ray mammography) through to surgery (surgical planning and image assisted surgery) and into therapy (e.g. assessment of therapy). Figure 4 shows a tree diagram of medical image analysis applications, focusing on breast-cancer imaging and showing the different modalities involved in clinical management. For each of these modalities research aims to develop computer applications that can assist the clinician to understand the underlying pathology. As shown in the diagram, our work is focussed in two areas that seem certain to play a role in the computer-assisted interpretation of digital mammograms.

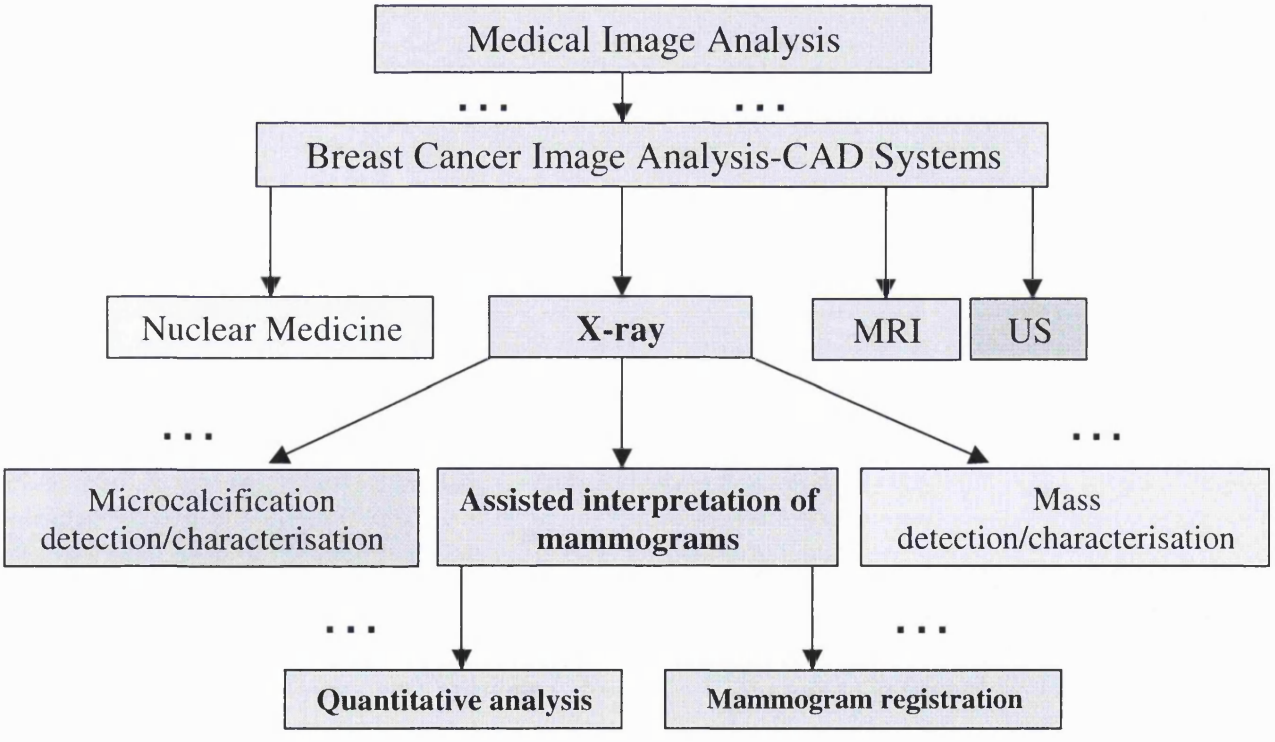


Figure 4: The different modalities involved in breast-cancer image analysis. Our work is focussed on applications relating to the assisted interpretation of mammograms.

In the remainder of this chapter, the more important modalities for breast cancer imaging are briefly presented and compared to X-ray mammography that is more important for the work presented in this thesis. In addition, the areas of development of CAD (Computer Assisted Diagnosis) systems for mammography are discussed. Finally, the possible applications of the presented work are considered in the context of medical image analysis.

2.2 Image modalities for breast cancer

2.2.1 X-ray mammography

Introduction

A mammogram is a specialised x-ray image of the breast. Mammography is the oldest but still the most valuable technique for breast imaging. The historical evolution of mammography can be summarised as follows:

- 1913: The first work on mammary roentgenography is reported [8].
- 1956: Large-scale mammographic screening was proposed, aiming in the early detection of cancerous, non-palpable masses [2].
- Early 70s: The introduction of the screen-film system, replacing the direct exposure film [4,8]. The intensifying screen placed under the film reduced the radiation dose to the patient and the exposure time required.
- 2000: GE's full field digital scanner, *Senograph*™* gains approval from the US Food and Drug Administration (FDA), for the reading of mammograms printed on film [9].

To obtain a mammogram, the breast is compressed between two plates, situated between the x-ray source and the detector, which is usually a film. Inside the breast, each x-ray beam is attenuated according to Beers Law:

$$I = I_0 e^{-\int \mu dx}$$

* General Electric Medical Systems [9]

Where I_0 is the incident x-ray intensity, I is the resulting intensity of the x-ray beam exiting the tissue, d_x is the thickness of the tissue and μ is the linear attenuation coefficient of the tissue. The different types of tissue inside the breast have different attenuation coefficients. This is why they attenuate and absorb different proportions of the x-ray beams. These differences are finally translated into different intensities in the exposed film. In film-screen mammography the single emulsion film is in contact with an intensifying screen which acts as a photon-amplifier.

In order to reduce the radiation dosage and to produce a roughly even thickness, the breast is compressed between two plates. This way, image quality is improved and is (roughly) uniform across the mammogram. That happens because if the breast were simply placed between the source and the detector, according to Beers Law, variation in the attenuation thickness due to the shape of the breast would subsequently cause a variation in the quality of the resulting image. The breast tissue distortion due to compression is not considered to produce artefacts in the resulting image. Usually two different views of each breast are obtained a cranio-caudal (CC) and a 45° medio-lateral (ML), in order to improve the diagnostic value of mammography.

One other interaction between x-rays and tissue is an example of Compton scattering, where the path of the photons is altered after collision with breast tissue. Such photons do not follow the path perpendicular to the detector, hence they cause image distortion. This phenomenon is restricted in practice by placing a grid before the detector, in order to stop photons that are not perpendicular to the detector, reaching the detector. However, this results in a substantially increased radiation dose to the patient. The anti-scatter grid can be modelled and removed using the h_{int} representation of interesting tissue (described in the Appendix C of the thesis) [4].

Strengths and weaknesses

The main advantages of X-ray mammography are:

- Good signal-to-noise ratio and high spatial resolution (~0.1 mm) resulting in a detailed representation of breast anatomy.

- Affordable cost, so that it can be used as the main screening technique and to be widely available in hospitals

- Good sensitivity (in [10], is reported to be 90%), for postmenopausal women.

- Its effective for microcalcification detection and pathology (in particular Ductal carcinoma in situ)

Although mammography remains the most reliable technique for diagnosing breast cancer, it has a number of significant limitations:

- The woman receives a radiation dose each time she has a mammogram.

- Breast compression cannot be tolerated by all patients. It also results in a non-rigid tissue motion, making the reading and temporal comparison of mammograms harder.

- The dense glandular tissue in the breasts of younger women results in bright, low contrast mammograms, thus making detection of masses a difficult task and imposing an age constraint on the use of mammography, as has been shown to be diagnostically useful only to post-menopausal women.

- It is obvious that the intensity of the x-rays reaching the film depends on the path taken through the breast and the resulting image is a 2-D representation of the 3D compressed breast tissue. This is a significant drawback of x-ray breast imaging since important tissue features may overlap or be occluded. In addition, surgeons need a 3D representation of the breast for guidance and surgical planning.

- Since imaging anatomical information is taken at a single instance, mammography cannot provide functional information (e.g. tumour angiogenesis, activity or increased blood flow). In addition, mammography is not very effective in evaluating the effectiveness of chemotherapy, since the adenosis and neovascularity around the tumour cannot be distinguished clearly by examining the film.

- Besides dense tissue, it can be non-diagnostic (or equivocal) in women with breast-implants, scar tissue (can appear as a cancer), lumpy breasts (with diffuse areas of increased or decreased density) and for multi-focal or multi-centric lesions where the extent of the cancer may be underestimated. In addition, the specificity in differentiating benign to malignant masses has been reported to be 20-50% [11,12], which reflects the high number of false positives.

Nevertheless, to date, film-screen mammography is the most important technique used in clinical practice. Recently, it was reported that mammography currently offers the best performance with respect to cost [13].

2.2.2 MRI imaging of the breast

Introduction

Nuclear Magnetic Resonance imaging is based on the magnetisation of human tissue. This is done first by applying an external magnetic field and then an additional RF pulse in order to displace and measure the tissue magnetisation $\boldsymbol{M}$ using an appropriate coil. Translating signals that come from different tissues into intensities at the corresponding points produces the MR image. The strength of each signal and subsequently the contrast is proportional to the magnetisation of each tissue in the xy plane (transverse magnetisation) [14,15]. In particular, the intensity of a pixel in a grey-level MR image is directly proportional to the strength of the signal emitted from a tissue voxel [15]. Subcutaneous fat emits a particular strong signal, and appears very bright in the MR image, while muscle emits intermediate signal and air in the stomach negligible signal.

Various Nuclear Magnetic Resonance (NMR) signals can be generated by manipulating $\boldsymbol{M}$ with suitable RF pulses (e.g. spin-echo pulse sequence [15]). Despite the continuous efforts, a pulse sequence that can distinguish between healthy tissue and tumour has not been developed yet. For this reason, is more effective to use MRI in combination with a contrast agent. In that way the dynamic behaviour of different structures within the breast can be monitored (functional imaging). In particular malignant tumours exhibit an increased vascularity, since they begin to grow their own blood supply network [14]. For that reason when the contrast agent is distributed, malignant masses enhance faster. This is illustrated in Figure 5.

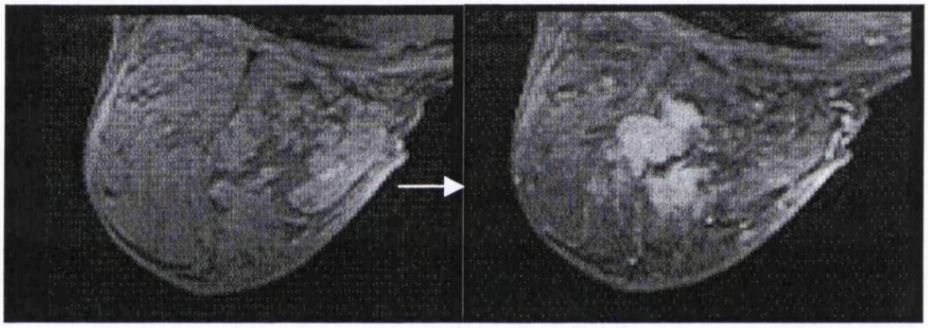
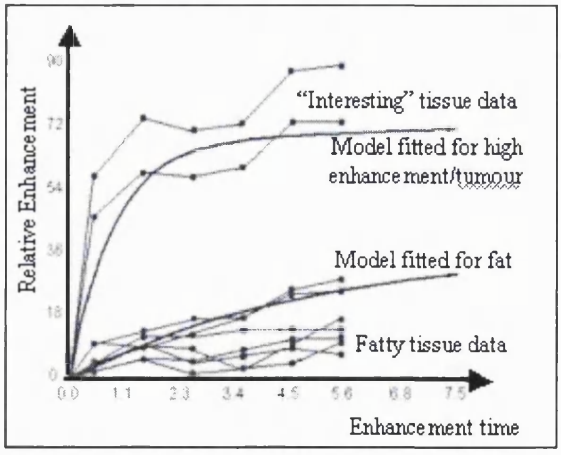


Figure 5: Enhancement in breast MRI by using contrast dye (images taken from [19])



(b)

Figure 6: The pharmacokinetic model separates high enhancement tissue structures (e.g. tumours) from low enhancement voxels (e.g. corresponding to fat).

Enhancement models have been introduced in order to detect regions of high enhancement such as tumours. Figure 6, shows the model fitted for high and low enhancement data respectively [17,18]. During image acquisition, such a model of relative enhancement can assist the clinician to quickly detect an abnormality, since the enhancement curves are significantly different in the first minutes of the acquisition.

Strengths and weaknesses

There are some significant trade-offs in MRI imaging. By increasing the voxel size, the emitted signal is increased and subsequently the signal to noise ratio of the image. If the cross-section of the voxel is increased, the trade-off will be a decrease in the spatial resolution of the image. On the other hand, by increasing the voxel thickness (or equivalently the slice of tissue being imaged) information concerning different tissue borders is lost and problems encountered with partial volume averaging become worse [15].

In clinical practice, the main advantages of the technique are:

- Unlike mammography, there is no harmful radiation and, by using a contrast agent (e.g. Gd- DTPA*), tumour vascularity can be imaged thus enabling the radiologist to differentiate between benign and malignant lesions.

- The intrinsically3D nature of MRI can assist the clinician in surgical planning and more general in minimal invasive surgery in the breast.

- It is suitable for younger women, where mammography is less effective for diagnosis.

- By taking images pre and post chemotherapy, one can assess the tissue response to chemotherapy.

* Gadolinium diethylenetriamine pentaacetic acid

- Offers good sensitivity and spatial resolution. However, a digitised mammogram at 50 microns ($50 \cdot 10^{-6}$m) has 400 times better resolution than an MRI scan of $1mm^3$ voxels.

- Suitable for imaging women with implants or scar tissue due to prior surgery.

On the other hand some weaknesses of breast-MRI include:

- Low specificity (~40-60%), since fibroadenomas, benign disease, areas of inflammation, and active glandular tissue can all show enhancement [20,21].

- Requires high infrastructure and the cost of MR scanners is very high. As a result, it is not so widely available as X-ray mammography is.

- The time of the acquisition remains long (15-20mins) resulting in patient motion that makes functional studies of the breast tissue more difficult (registration of the temporal sequences is required).

- Contrast agents are toxic. In fact 1 in 10^4 women has a side effect.

- Sensitivity to small lesions and DCIS (ductal carcinoma *in situ*) is poor to average but improving. In addition, microcalcifications are almost impossible to detect since clinical MRI is based on soft tissue.

2.2.3 Ultrasound

Introduction

Ultrasound is defined as an acoustic wave of a frequency above human hearing [15] (usually around 7-20MHz). In the early 1950s, it was demonstrated that ultrasound could image breast structures and distinguish cystic from solid masses. Currently, ultrahigh

frequency probes are available with superb axial and lateral resolution, and extended dynamic range. Ultrasound of the breast structures is acquired by recording their reflections of ultrasonic waves. For that reason, the technique is also known as pulse echo imaging. The reflected waves coming from different tissues and from different depths within the breast carry all the information necessary for the image to be formed.

Because of the way the ultrasound image is produced (wave propagation-detection) it can be considered a tomographic image of echoes, based on the echogenic variation among different structures (tissues). The skin and calcifications exhibit the strongest echoes, while dense glandular tissue and Cooper's ligaments are less echogenic. Finally fat tissue, benign tumours and cancers are even less echogenic and fluid is almost unechoic [22]. It is worth mentioning that the strength of the reflected wave (described by the reflection coefficient R) is dependent on the acoustic impedance of both tissues (Z_1 and Z_2) that form a reflecting boundary. Figure 7 ([23]), illustrates the appearance of an ultrasound scan with a cyst present.

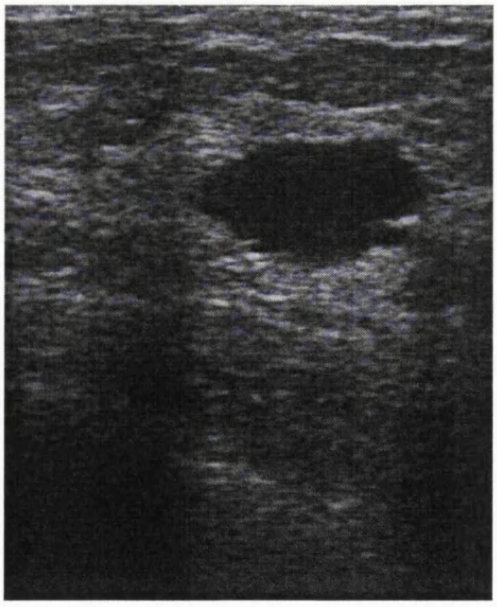


Figure 7: Appearance of ultrasound imaging of the breast. A cyst (unechoic dark region) is present.

Strengths and weaknesses

The main advantages of ultrasound are:

- It's inexpensive and real-time acquisition, widely available. It is also possible to obtain 3D ultrasound images, usually by using a free-hand probe followed by 3D reconstruction. Recently, contrast agents have become available for assessing tumour vascularity.

- Ultrasound can reduce the number of unnecessary biopsies and allows image-guided biopsy.

- It can detect cysts with high accuracy as well as macrocalcifications and it's suitable for women of all ages.

Some of the weaknesses include:

- Low sensitivity. Images are intrinsically very noisy and hard to interpret. However, new high frequency scanners exhibit improved spatial resolution (e.g. the GE LOGIQ™ ultrasound systems uses 12 MHz imaging, and has demonstrated point resolution under 300 microns, approaching that of CT and MR systems [9].

- Poor specificity in cancer detection (depends on the experience of the user). In [24], the sensitivity and specificity of sonography for cancer detection is reported to be 100% and 48% respectively.

- The breast is deformed during acquisition. Due to tissue deformation, it can become difficult to compare ultrasound information to other modalities.

2.2.4 Nuclear Medicine for the breast: PET and Scintimammography

Introduction

The application of nuclear medicine in breast imaging is very promising. In nuclear medicine, a radiopharmaceutical (a radiolabelled tracer attached to a metabolite) is injected into the patient. The metabolite binds preferentially in certain regions, and the radiotracer begins to emit particles that are picked up by a scintillator and the emitted radiation reveals information about a specific organ. Radiollabelled oestrogen and glucose analogue uptake is more intense at sites of tumour growth or lymph node metastasis. In particular the glucose analogue FDG (Fluorodeoxyglucose) is known to concentrate in breast tumours rendering them easily detectable in conventional PET scans [14, 25].

Besides PET, Scintimammography is becoming more popular, mainly because it is less expensive compared to PET. In scintimammography, ^{99m}Tc sestamibi is intravenously administered to the patient and an emission image is acquired 5-10 minutes after the injection using a gamma camera [26]. Since the uptake in tumour cells is nine times greater then in normal cells, the presence and location of a lesion is identified. In clinical practice either planar or SPECT imaging is used. In planar imaging, six different views are acquired (four in the prone position and two in the supine), while if SPECT is used image slices of the chest (through 360° around the patient) are gathered thus allowing the localisation of tumours in 3D. This way, scintimammography can be useful for surgical planning and biopsy-guidance.

Figure 8, shows the correct positioning of the patient in the supine position during acquisition, while Figure 9 shows some typical scintimammograms and Figure 10 the use of the technique for evaluating chemotherapy (images taken from [26]).

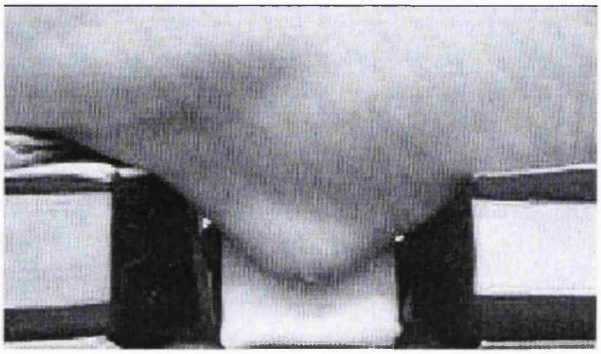
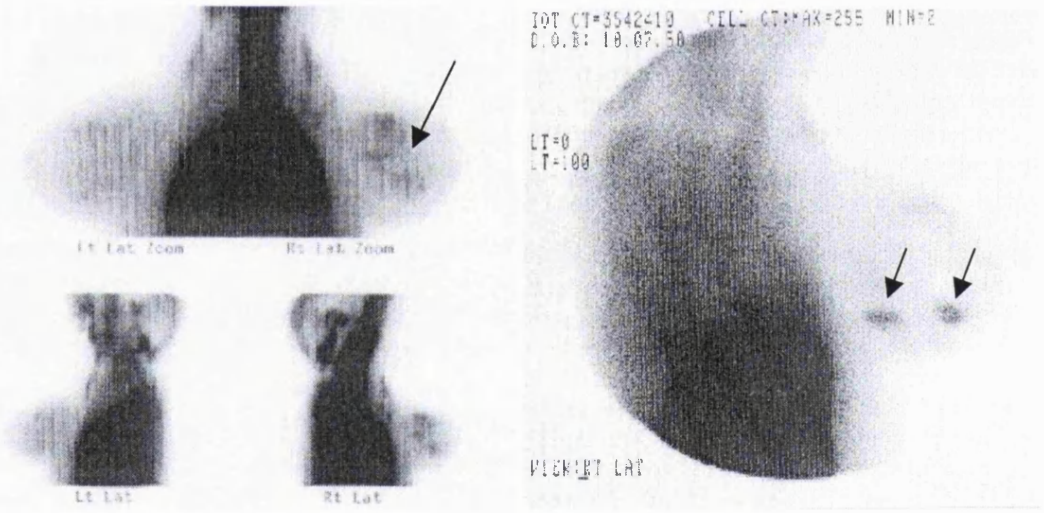
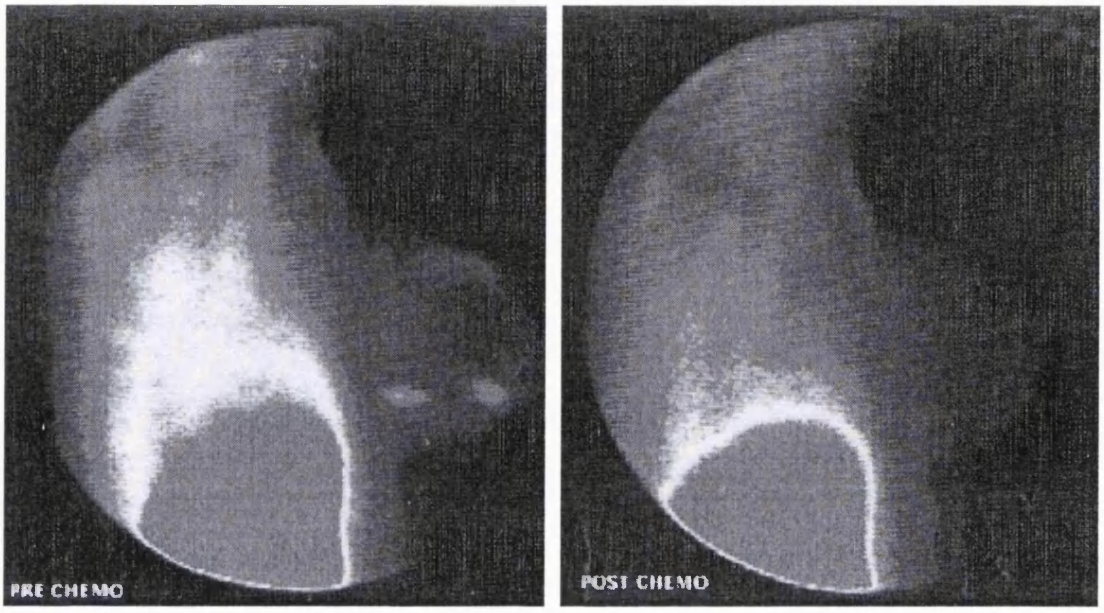


Figure 8: Correct patient positioning with the patient in the prone position and the breast unconstricted by the imaging couch and enlogated by gravity.



(a) (b)

Figure 9: Scintimammogram showing (a) multi-centric and (b) multi-focal cancers



(a) (b)

Figure 10: Using scintimammography for evaluating chemotherapy: Image (b) illustrates the reduced activity uptake in the tumour after treatment.

Strengths and weaknesses

The main strengths are:

- It's almost as sensitive as X-ray mammography and MRI in palpable tumours but with greater specificity [26].

- In addition, it can give significant diagnostic information in radio-opaque mammograms of dense tissue (e.g. younger women) that are difficult to interpret.

- Scintimammography is less expensive compared with MRI [26].

- It can provide functional information and thus be used for detecting cancerous masses (angiogenesis) and can monitor the effects of chemotherapy in patients diagnosed with breast cancer.

- It can effectively image the axilla, thus providing useful clinical information concerning breast cancer metastases. This is particularly useful as 45% of cancers appear in the upper outer quadrant (UOQ) of the breast (a fact that is further discussed in chapter 5).

The main weaknesses of nuclear medicine techniques for breast imaging are:

- PET imaging for the breast is very expensive and not widely available.

- For Scintimammography, the most significant drawback is its low sensitivity in lesions (<1cm) or in other words the low resolution images acquired.

2.2.5 Conclusion and comparison to X-ray mammography

In clinical practice, there is always a combination of imaging and clinical examination of the breast. To date, the most common combination is the triple assessment, which includes mammography (for women 35 and over) or ultrasonography (for women under 35), clinical examination and fine needle aspiration cytology. According to J.M. Dixon, in a recent investigation concerning 1511 patients with breast cancer, only in 6 patients all the three investigations of the triple assessment, turned out to be benign [27]. Although mammography is still the most widely used technique for breast-screening and minimal invasive surgery, new modalities have increasing promise, particularly in giving functional information about breast cancer:

Ultrasound has been used in combination with x-ray mammography for many years. In most cases, if the mammogram has a lumpy appearance or if it shows a well-defined mass, ultrasound examination is recommended in order to decide upon whether the lump is a cyst or not. Recently, the introduction of 3D ultrasound and contrast-enhanced ultrasound has shown some promising results, but their role in breast imaging is still to be demonstrated [26]. The cost of an ultrasound scanner is comparable to the cost of an X-

ray scanner (around $100,000 US) but the examination cost is higher since a well trained clinician (compared to a radiographer in mammography) is required. The acquisition time for a 3D freehand ultrasound scan is around 5 minutes (comparing to 1 minute for the mammogram acquisition and 10-15mins for MRI and scintimammography).

Breast Magnetic Resonance Imaging is used today as a supplement to, or sometimes in combination with x-ray mammography. It is suitable for monitoring younger women and women who have previously had surgery (as scar tissue can produce artefacts in x-ray mammography). Several studies have proved that although it has high sensitivity for detecting masses, its specificity is often very low [14]. Nevertheless, breast MRI is a useful technique, especially for younger women with dense glandular tissue. Contrast-enhanced MRI can detect small enhancing lesions: in [28], 3mm (diameter) DCIS were identified in four patients while the best images can show a resolution of about 2mm [29]. In another study [30], dynamic contrast MRI was found to be superior to triple assessment for multi-focal disease detection (with sensitivity and specificity 73%, 100% for MRI and 18%, 100% for triple assessment). However, the significant cost of MR scanners (typically an order of magnitude higher than a film-based X-ray scanner) as well as the infrastructure required (special room, shielding, specially trained staff) limits its use.

Nuclear medicine imaging of the breast (e.g. Scintimammography), can also provide functional information as a cancer takes up the metabolite more intensively than the rest of the breast tissue, without any age or surgical history constraints. Compared to mammography, the higher cost and the low availability are the main reasons for the limited use of this technique to date. However, it can provide functional information and can be used for imaging younger women (especially those in high-risk groups). Compared to MRI, scintimammography has lower spatial resolution (MRI sensitivity is nearly 100% exhibiting excellent anatomical information), but higher sensitivity while being less expensive [26]. It has been reported that tumours less than about 10mm in diameter are difficult detect [14]. However, new scintillation cameras promise increased spatial resolution (e.g. in [31] the spatial resolution is around 5mm).

Table 1, summarises the main advantages of X-ray mammography as well as the need to use and further develop other modalities in clinical practice. Evidently, it is desirable to be able to combine information from different modalities in order to diagnose pathology more reliably. Our work in mammogram registration is a key step towards an imaging system that enables temporal and multi-modal fusion for asymptomatic women (that are enrolled in the screening programme) and for women that present to the clinician with a lump.

As is discussed in chapter 6, the mammogram registration work, presented in chapter 4, has been extended to multi-modal registration between X-ray mammography and breast-MRI [32, 33]. We believe that temporal registration and data-fusion will play a key role in future diagnostic computer systems, enabling the alignment of images from several modalities and comparison of patient data with previous ones. Temporal registration can be considered as a basic step towards a multi-modal fusion system for breast cancer imaging. The next section summarises the main research areas in CAD (Computer Assisted Diagnosis or Detection) systems for mammography and discusses the potential of our research in clinical practice.

Advantages of X-ray mammography:	**Need for other modalities:**
• Good anatomical representation • Offers the best performance/cost • Is used in the screening programme	• Where mammography is equivocal or non-diagnostic • Younger women/dense tissue • Differentiating benign/malignant • Surgical planning/Minimally invasive surgery and chemotherapy evaluation

Table 1: The main benefits of mammography vs. the rationale for using alternative modalities in clinical practice. Instead of having to choose the best modality, it is more reasonable to combine multi-modal information (although is a difficult engineering problem) through *data-fusion*.

2.3 Digital mammography and Computer aided diagnosis (CAD) systems

As illustrated in Figure 1, CAD systems for breast-cancer, aim to assist the clinician to interpret images and establish reliable and early diagnosis of pathology. Digital mammography will increasingly replace conventional film-screen mammography systems even if at the moment the cost is very high compared to the film-based device ($400 000 - $600 000 compared to $50 000 - $70 000) [13]. It has been reported that it offers better contrast, dynamic range and improved diagnosis rate of masses [6]. In addition, it captures images directly and stores them digitally, facilitating image transfer between clinicians and safe storage with rapid recall. The main advantage, from the image analysis point of view, will be the direct integration of CAD systems that will enable the clinician to manipulate the mammogram while he or she is reading it (e.g. image enhancement, registration with previous scans and automatic detection of features). However, the transition from the film-screen environment (where the film is read by placing the film onto the lightbox) to the digital one (where the clinician will read the mammograms from the computer screen) is not trivial. The SCREEN project [7], is a European effort to assess the difficulties in this transition and suggest solutions for efficient hardware implementations of digital-reading systems.

Many different algorithms for aided detection of mammographic features had previously been suggested. Moreover, the first CAD system to achieve FDA approval was the R2 Image checker [34]. The main areas of research related to mammographic CAD systems are:

Microcalcification detection and characterisation. Since microcalcification detection can improve the early diagnosis of subtle cancers, many algorithms have been suggested for detecting clusters. In addition, several algorithms have been developed to try and characterise clusters as benign or malignant, based on the geometry of the clusters and quantitative measurements. It has been reported that using the *R2 ImageChecker®*, microcalcifications are detected with more than 98% of accuracy. However, as is reported in [35], although the diagnostic sensitivity of the clinicians rose when using the system,

the positive predictive value of the clinician's interpretations worsened due to the high number of false positives. Recently, research in 3D-reconstruction of clusters using two views (CC and ML) has shown very promising results [36, 37].

Mass detection characterisation: Such algorithms aim to automatically characterise cancerous tissue against normal parenchymal tissue and benign disease. Characterisation as benign or malignant is usually done either by describing the shape of a lesion (e.g. circumscribed or speculated) [38], or by characterising the texture calculated in the region of the mass (e.g. using fractal analysis measures followed by the design of a classifier [39]). Using the R2 system, the accuracy in detection rate was reported to be less than 73% for mass detection [34].

Computer –assisted interpretation of mammograms: This is a more general category that includes image analysis techniques that allow the clinician to better understand the mammographic features. In addition, there is an effort to combine image information derived from CAD algorithms with protocol-based aids to patient management, in order to assist the clinician in the decision making [40]. These ideas have inspired the development of decision support systems for mammogram interpretation [41, 42].

Recent advances in medical image analysis have shown potential to influence the interpretation of radiological images. A good example is the curvilinear structure (CLS) detector, where structures such as vessels are identified and can be removed, in order to facilitate the reading of the mammogram. This is because removal of the CLS from an image enables better detection of tumours and in itself can provide useful information about the location of objects such as calcification [43]. The work presented in the thesis is focused on two important aspects of such interpretation systems:

- Mammogram registration: Mammogram images have very variable quality and anatomical content. Hence, diagnosis from a single mammogram is intrinsically difficult. For this reason, if previous mammograms are available, radiologists often use two (or more) such and compare /contrast them to detect significant change(s). A

major problem is the geometric mis-alignment and difference in intensities because of differences in image formation. The work presented in chapter 4 aims to make that comparison more efficient and significant through temporal mammogram alignment.

- Quantitative comparison: When tissue density changes are of significant interest (e.g. in the case of women using HRT), it would be very useful for the clinician to have quantitative information concerning tissue changes. In the case of HRT sequences, it is important to follow the evolution of glandular tissue content after the beginning of therapy in order to assess the response to therapy and the risk of a hormone-stimulated abnormality. To date, quantitative comparison of tissue density over time using mammography, is not possible. The work presented in chapter 5 aims at the development of measures of tissue density change that could enable the quantitative comparison of HRT mammogram sequences.

- Other CAD applications include:
    - More generic image processing and manipulation algorithms (e.g. Image enhancement using wavelet filtering)
    - Detection of radiological tags, or anatomical points in mammograms
    - Noise removal, breast edge enhancement.

2.4 Conclusions

Although new modalities have shown many promising results for the early diagnosis of breast cancer, X-ray mammography is still the most widely available. Data fusion between mammography and newer modalities is the most reasonable choice in order to exploit the medical information available from different modalities. CAD systems for mammography aim to assist the clinician in the interpretation of mammograms and in diagnosis. In addition, the large amount of data arising from subsequent visits to the screening programme, encourage the development of computer applications that can assist the clinicians (and especially the less experienced ones) in the reliable early diagnosis of cancer.

Our work is a result of a conscious choice of clinical problems and aims to facilitate temporal mammogram comparison and provide quantitative measures of tissue density-change for women on HRT. Temporal registration and quantitative comparison, could be important for:

1. Comparison to previous mammograms in subsequent visits to the screening programme. This way, it can become easier to detect changes such as benign and malignant masses, asymmetry and calcification clusters.

2. Interval cancers: As was mentioned in the previous chapter, it is believed that a significant number of interval cancers are "missed" in previous readings. A more effective temporal comparison (e.g. computer-assisted comparison using registration) could increase the detection rate of cancers thus reducing the "missed" interval cancers which represent a basic weakness of the screening programme.

3. HRT users: As will be discussed in chapter 5, there is evidence that substantial localised changes in tissue density for women using HRT (especially long term users) seem to be correlated with cancer. Our work aims to assist in the clinical assessment of the woman's response to HRT (through quantitative comparison with previous mammograms) and in the case where the woman responds to therapy (which is often associated with an increase in glandular tissue density) register future mammograms to the new baseline (the "increased density/ changed pattern" mammogram) in order to detect an abnormality early.

4. Retrospective studies to assess the performance of radiologists in the screening programme. In other words, cancers that have been missed because the clinician failed to detect signs of abnormality (e.g. very subtle mass, architectural distortions, ill-defined masses and focal symmetric densities). The mammogram in which the cancer was diagnosed is registered to previous ones so that it can be

assessed if there was enough information to diagnose the disease at an earlier stage.

5. Registration could be used to educate radiologists since it becomes easier to follow the evolution of regions of tissue over time through physiological (e.g. involution) or pathological changes in aligned sequences.

Summarising, computer aided temporal mammogram analysis has the potential to assist the clinician in the detection of subtle cancers that as mentioned are often missed. Establishing correspondences between temporal/bilateral mammograms could be an important aspect of CAD (computer aided diagnosis) systems. Registration (alignment) of mammogram sequences can assist the clinician to better understand the correspondences between temporal mammograms and diagnose abnormalities. Furthermore, quantification of tissue density changes can assist the clinician to understand if the woman is at higher risk of cancer due to HRT. Our clinical co-workers believe that the development of robust temporal registration and quantification systems could reduce the number of "missed" cancers and assist in the early detection of cancers related to the use of HRT. Future clinical trials (e.g. in the screening programme) will test these expectations in clinical practice.

Chapter 3: Review of Medical Image Registration

3.1 Introduction

In this chapter the basic concepts of medical image registration are explained with focus on mammogram registration, which is the main subject of our work. In clinical practice, medical imaging has traditionally been used for diagnosis. Over the past few years, a number of medical imaging applications including surgical planning, image-assisted surgery, radiotherapy and chemotherapy assessment have become increasingly clinically important.

A coarse classification of medical imaging can be made in terms of anatomical and functional imaging. Anatomical imaging modalities, such as X-ray, CT (computed tomography), MRI and ultrasound provide information about the shape and structure of various parts of the human body. On the other hand, functional modalities like contrast-enhanced and fMRI (functional MRI), nuclear medicine (e.g. scintimammography), SPECT and PET provide information about changes in physiological processes (e.g. contrast-enhanced MRI provides information about the breast tissue uptake of the contrast agent over time).

Different modalities have their advantages and disadvantages and it is most often the case that the clinician needs to combine information from different modalities, or of the same modality taken at different time instances. The large variability of medical image data and the diverse and often complementary nature of the images, lead to the development of techniques that would allow the fusion of information coming from different imaging modalities. This process is generally termed *registration* and covers a wide range of applications and modalities.

In this chapter, the different aspects relating to medical image registration are explained in more detail.

3.2 Definition of concepts related to registration

The terms *registration*, *fusion* (or image-fusion) and *matching* are often used to describe similar concepts (and sometimes can cause confusion to the reader). In this section, some short definitions are provided in order to help the reader understand the use of these terms in the rest of the thesis as well as some comments about the image transformation aspect of the registration process.

Image Matching: This is defined as the process in which correspondences are established between two or more images using an algorithm that calculates similarities between structures (primitives) in those images. These structures can either be smaller anatomical elements (in which case a segmentation algorithm prior to matching is necessary) or characteristic/salient points, lines, contours or edges (usually texture masks, edge/line detectors are used to define/segment them). The correspondences can be calculated for all the pixels in the image (as, for example, in an optical flow algorithm) or between a set of image primitives like patches (in general using a similarity measure like mutual information or texture in general). Depending on the individual modality and the images, a set of constraints or heuristic observations can be defined in order to help establish correspondences.

Image Registration: As noted in the introduction to this chapter, it is most often the case that the clinician needs to combine medical images from different modalities, or images taken at different times. The first step in this integration process is the geometrical alignment of the images. This process generally involves estimation of the transformation needed to align an m-dimensional image with an n-dimensional image (n, m $\in$ 2,3,4). The dimensionality of the registration is discussed later in this chapter. Matching is the first step of a registration process, since it is necessary to establish some correspondences in order to drive a spatial transformation to align the images.

Image fusion: After geometric alignment the next step of integrating information from different medical images, is image fusion. It is defined as the integrated display of the

data involved [44] (e.g. a cancer detected in a nuclear medicine scan, superimposed on a geometrically aligned 3D MRI scan of the same patient). In some applications, image fusion is the most important part of the registration process, the finest example being pre-operative planning augmented by intra-operative imaging, or fusion of CT with the patient's video images [45, 46].

Comments on the registration process: The number and type of correspondences between images clearly defines the choice of registration technique, the mathematical extent of transformation (i.e. rigid/affine/non-rigid) and the type of heuristics/constraints necessary to make the registration robust. Additionally, there are a number of optimisation considerations that must be taken into account in order to ensure that the choice of correspondences produces a computed transformation that correctly compensates for the differences in the images that need to be registered. Finally, the optimisation framework must reflect the type of regularisation (or constraints) imposed on the registration process.

In terms of computing the transformation, there is a large degree of flexibility in the order of the computation. The main classifications are:

- Rigid: When the necessary transformation only needs to translate and rotate one image to be aligned to the other(s) (3 degrees of freedom).
- Affine: Same as rigid but in addition a scaling parameter is necessary for the alignment (4 degrees of freedom).
- Locally Affine: This method has been used for surface matching [47], combining an initial global affine transformation, then a local affine modification to complete registration
- General non-rigid: When more parameters are necessary in order to align the images, the transformation is non-rigid (high-order degrees of freedom).

The transformation that relates the two images depends on:

- The difference in geometry arising from the way an organ is imaged in each modality (e.g. in X-ray mammography the breast is compressed between two plates, whereas in breast MRI it is uncompressed and pendulous).
- The patient positioning: For example a mammogram is taken while the patient is standing up, whereas in the breast-MRI acquisition the patient lies in the supine position.
- The intrinsic differences (e.g. the image formation process, or dimensionality) between modalities can introduce geometrical inconsistencies (e.g. breast MRI and nuclear medicine can both provide functional information, but nuclear medicine provides limited anatomical details).

In most cases, the clinical application (i.e. the characteristics of the organ or region to be registered) will determine which type of transformation is appropriate. For example, soft tissue medical imaging (perhaps with the exception of the brain) will almost always produce a non-rigid transformation between successive acquisitions due to factors such as patient movement, muscle flexing, breathing, cardiac motion, etc. In many such cases, there is a need first to compute a general transformation, perhaps based on a rigid or rigid/affine transformation, and then to proceed to refine the transformation on the basis of a non-rigid relationship between corresponding points.

Alternatively, the type of transformation may be driven by the characteristics of the correspondence data. For example, when registering brain MRI images to an atlas, in many cases, a rigid transformation may be suitable. However, if the correspondence data indicates a higher-order transformation (e.g. affine) on the basis of some minimisation framework, then the transformation could be adjusted to reflect a higher degree of complexity. This also applies to applications where the transformation may be predominantly rigid or rigid/affine, but has a small non-rigid component that may vary from case to case (e.g. sequences of breast images).

These considerations are addressed more thoroughly in section 3.7 of this chapter.

3.3 Clinical Motivation for Registration

There is a wide range of clinical applications that have motivated the need for computing correspondences and the transformation characteristics between images. Although several techniques allow automatic registration using stereotactic frames, fiducials or skin markers, most of these scenarios have limited applications (e.g. it is not possible to have markers in all modalities) and their usually invasive nature make their use less attractive. There is a need for frameless guidance systems to help surgeons plan the exact locations of incisions, to define the margins of tumours and to precisely identify locations of neighbouring critical structures.

In general, the motivation for developing sophisticated registration techniques has been on the basis that there is no single defined characteristic in the image that can be used for correspondence. As a consequence, nearly all registration algorithms involve developing correspondences between detected features such as points, landmarks or regions of interest. The applications of registration (to date) have been designed to address the following types of clinical applications:

- Diagnosis, screening and temporal studies. Registration can assist the clinician in establishing a diagnosis via a more efficient comparison of multi-modal or temporal mono-modal patient data. Screening for a disease (e.g. for breast cancer) provides the opportunity to compare previous data of the patient with the most recent, and detect abnormalities on the basis of significant difference (as in the case of temporal mammography). In addition, image registration can provide useful information about the progression of a disease or treatment since the aligned images can show the differences between the images over time (e.g. in HRT, chemotherapy, radiotherapy).

- Fusion between different imaging modalities in order to combine information from the two images in a single image. This is a general clinical requirement with

several applications. As an example, recent work on data fusion between X-ray and breast MRI has shown promising results [32, 33].

- Generation of statistical models for comparing "normal" against pathological conditions. To build such models, it is necessary to acquire information from a large number of controls (i.e. healthy individuals) and examine the "normal" variation of certain features (size/volume, texture and shape). Image registration is needed in this step in order to align all the images to one co-ordinate frame. Once the model is computed a patient can be diagnosed "abnormal" if the data does not fall within an allowable variation. The basic assumption behind models of normal variation is that there is a consistent difference in features (e.g. volume or size of a structure) between "normals" and patients with a disease (e.g. in schizophrenia). In [48], a probablistic atlas describing the normal variation of the human brain is constructed based on high-dimensional vector field transformations, while in [49] a biomechanical model (that can simulate tumour growth) is used to register brain images to a "normal atlas" and accurately detect abnormalities for effective surgical planning.

- Image guided surgery. After image acquisition and surgical planning (including segmentation and/or simulation), the medical image and planning information is registered to the patient in the operating room (OR). The surgical planning may involve registration of a pre-operative model to the intra-operative conditions of the OR. In the same context, surgical guidance (fusion between intra-operative geometrical information and a pre-operative model) can assist (navigate) the surgeon during the procedure [50]. Such methods enable a visual mix of live video of the patient with the segmented 3D MRI or CT model, supporting enhanced reality techniques for planning and guiding surgical procedures, and to interactively view extracranial or intracranial structures non-intrusively [51-53]. Extensions of these methods include simulation of tissue change during surgery [54-56], image-guided biopsies [57] and focused therapeutic procedures.

- Atlas or template matching – for quantification studies in cardiac imaging [58], or for comparing a known library of anatomical/physiological information (this can also be functional, in the case of fMRI, for example) against a specific patient case [59, 60]. Usually a deformable model is used to transform the individual images to the co-ordinates of the atlas for easier and more consistent comparison of patient data [61].

3.4 Specific Motivation for Registration in this Thesis

In this thesis, the major emphasis is on registration problems related to mammography. As mentioned in the first chapter, comparison of mammograms of the same patient taken at different instances can assist the clinician in detecting pathological changes and hence establish a diagnosis of cancer at an earlier stage. As a consequence of the implementation of screening programmes, temporal mammographic data has become increasingly available and clinicians use temporal comparison as a way of detecting abnormalities. Registration of temporal mammograms can enhance this comparison and assist the clinician in detecting asymmetries, dissimilarities and new growths.

The key objective is to register (align) temporal mammogram images so that a clinical comparison becomes an easier task and the architectural similarity (or dissimilarity in cases where there is a new growth) is maximised. The method developed, and presented in this thesis can extract consistent geometrical landmarks along the edge of the breast in order first to align the images based on the boundary of the breast. In a second step, a multi-scale segmentation based on wavelet analysis provides a set of internal structures that can be matched in the two breast images and then refine the registration to better approximate the transformation. In this chapter, the nature of the mammogram registration problem is discussed in section 3.10 and in chapter 4 the technique for mammograms registration is presented.

A special case of mammogram registration involves temporal data from women that have been treated with Hormone Replacement Therapy (HRT). Since it has been reported that

there is a greater risk of breast cancer for HRT users, possibly induced by the exogenous hormonal stimulation, frequent mammogram screening is necessary for early diagnosis of pathological changes in breast tissue (especially for long-term users). In chapter 5 of the thesis several issues of temporal comparison and tissue quantification in temporal HRT mammograms are detailed.

3.5 Basis for Registration

In section 3.2 it was briefly mentioned that the objective is to use the image content to drive registration via the detection/computation of a set of correspondences. In general, the registration problem can be defined in terms of two categories [44]:

- Intrinsic, using image content such as landmarks (anatomical and geometrical), segmented features (rigid models based on points, curves, surfaces and deformable models like snakes, nets etc.) and voxel properties

- Extrinsic, using external markers, fiducials markers etc. These methods, though more accurate, are less attractive due to their usually invasive nature.

The Extrinsic case (which often includes non-image considerations such as the geometrical correspondence between images acquired in different scanners, etc.) is not the emphasis of this literature review. Since for breast registration one may not have the benefit of any obvious anatomical landmarks, there is a strong dependence on the computation of image-based features and characteristics to drive the registration process.

Shape constraints, topological models, snakes, adaptive meshes may provide some constraints or a "model" by which registration can be controlled (e.g. when registration is based on the evolution of contours or surfaces). In this case, a shape model is extremely useful for computing a point correspondence of the evolution of a contour (or surface). Shape-based approaches can also be considered either as an optimisation step, a model-based constraint (such as allowable shape deformations with respect to an atlas) or in

some instances, a regularisation consideration (for example, a contour may only be allowed to deform with respect to a certain degree of curvature, etc.). Useful information on various deformable models used in medical image analysis, can be found in [62, 63].

3.6 Types of registration

In the literature, the registration process is often classified according to the dimensionality of the problem, the modalities involved and the subject.

There are many different combinations of dimensionality in registration applications as for example 2D, 3D, 2D-3D (via projections), temporal 2D and 3D (effectively 4D as in [64-68]). The key requirement is that a geometrical framework can map one set of image feature vectors onto another.

Depending on the modalities involved, the registration problem can be characterised as:

- Mono-modal, if the images to be registered come from the same modality. For example in [69], where time series of MRI volumes are non-rigidly registered. Another example of mono-modal registration is temporal mammography where the most recent mammogram is aligned with all previous ones of the same patient, in order to detect abnormalities on the basis of significant difference. Another application of mono-modal temporal registration is to assess the success of the surgery (or therapy) by aligning the images before and after surgery (or the commencement of therapy).

- Multi-modal: When two or more different modalities are involved. In this case, it is necessary to register the different modality images in order to combine the information in a single display (data fusion). A classic case is the registration of PET to MRI in order to relate an area of abnormality to the underlying anatomy [70]. In this case, multimodal registration is necessary, since nuclear medicine is very sensitive at detecting lesions but has very poor spatial resolution, whereas MRI offers

excellent spatial resolution and anatomical detail. Another important example is MR-CT multi-modal registration, using for example a similarity measure [71, 72].

A multi-modal registration problem can be decomposed in two parts: geometrical alignment and intensity alignment. In this context, if the intensities of two images are corrected, the remaining problem can be treated as a mono-modal registration. This concept was exploited in [73], in order to take advantage of the demons algorithm (that was originally developed for the mono-modal case) in registering brain images from different modalities. In this thesis, the problem of mammogram registration though a mono-modality problem, can also be seen as a geometrical and intensity alignment (geometrical alignment via boundary and internal landmarks and intensity normalisation via the h_{int} representation).

If the images to be registered are acquired from a single patient, the registration is referred to as intrasubject. If different patients are involved the process is called intersubject registration. This is usually performed when it is desired to derive image-based statistical information about a disease (or anatomical, physiological information) from a population. For example, if a region of interest is to be compared between a normal (control) group and a group that has developed a certain disease (e.g. cancer), it is necessary to register the images to the same geometrical frame in order to facilitate computer calculations and make the whole process more robust. Ever since intersubject registration became possible, there has been an effort to gather information about normal variations to anatomical features. Statistical models of variation associated with a disease were discussed in section 3.3.

3.7 Analysis of the Registration Process

3.7.1 Overview

The technical aspects of the registration process can be decomposed into three distinct components:

- Correspondence of features or "matching stage"
- The computation of a transformation based on this correspondence
- Regularisation, optimisation of the registration
- Interactivity

These issues are discussed in the following sections.

3.7.2 Establishing correspondences

As mentioned earlier, the term registration is often used interchangeably with the term "matching". In fact, the matching process is really the first step of registration and involves determining correspondences between images so that a transformation can be computed. There are a number of ways to define correspondences that are usually directly related either to the imaging modality, the type of organ/region under examination or the clinical assessment that may result from registration (for example, if one is comparing schizophrenic brain against a "normal" population, contours are more critical than points).

In the literature, there are many different kinds of correspondence approaches:

- Points (i.e. salient regions, landmarks representing a region of interest, a collection of contour points, etc.). Various kinds of search algorithms (like the ICP algorithm described in section 3.8) are usually combined with local feature detection (corners, scale-saliency, scale-space analysis, etc.).

- Regions. Correspondence techniques such as block matching or "patch matching" use some kind of correlation measure or texture analysis to ascertain that there is a regional similarly between parts of an image. In general, this involves collecting a large number of "icons" within the image and utilising a cost function to compare the quality of similarity. Interestingly, most of the search algorithms used to define region

correspondence are similar to those for point correspondence. Similarity measures, and in particular mutual information, are described in section 3.8.

- Field (total image) Correspondence. These approaches attempt to compare every component in the images to be registered with the objective of defining a "transformation map" that defines how each local component differs between images. Clearly this approach is not generally applicable to simple rigid/affine registration applications as such correspondence measures generally indicate a large degree of local deformation. Examples of techniques that compute a transformation "field" are optical flow, fluid-based techniques and demons (discussed in section 3.8).

3.7.3 Transformation

Once a set of correspondences has been determined, this relationship is used to compute a transformation in order to geometrically align the images. Clearly, the number of corresponding points and the spatial relationship between these points will determine the class of transformation. As mentioned in section 3.2, the transformation calculation can be defined by a limited set of classes that depends on the clinical application. These classes are:

- Rigid: A co-ordinate transformation is called rigid when only translations and rotations are allowed
- Affine: If the transformation maps parallel lines onto parallel lines is called affine.
- Projective: The transformation maps lines to lines.
- Curved: Maps lines onto curves.
- Non-rigid: When none of the above restrictions can satisfactorily describe the transformation and higher degrees of freedom are required to "capture" the geometry.

In this thesis, the non-rigid case is more important, although in many applications the computation of a non-rigid transformation includes a substantial rigid or affine

component. This is clearly a consideration when computing the transformation and is further discussed in the context of optimisation in section 3.7.5.

Computation of the transformation for the rigid case is usually quite straightforward as only a limited number of corresponding points (e.g. only 3 points in 3D) is required to define a rigid transformation and calculation of the homogeneous matrix defining the transformation relationship can be performed quite robustly using simple gradient descent, least squares, convex matrix inversion, etc.

Non-rigid transformations are more difficult to calculate. Firstly, there is the issue of how well the set of corresponding points can be used to define a non-rigid transformation. If the data is sparse, computing the transformation can be numerically problematic. Additionally, there is the issue of whether to interpolate the transformation between a set of corresponding points or to use an approximation framework. This issue is discussed in the context of regularisation in section 3.7.4. There are many algorithms in the literature that have been used to compute a non-rigid transformation. These include:

- Various kinds of radial basis functions, including Gaussian [74], multiquadrics [75] and thin-plate splines [76].

- Different kinds of spline meshes and tensor splines (e.g. in [77] where vector splines are used for the reconstruction of the left ventricle. This is also extended to spline surfaces and spline mesh topologies.

- Constrained models such as finite-elements and mesh-based transformations (these are model-based deformation approaches like in [78]).

3.7.4 Regularisation

Despite the fact that a limited number of correspondences may be used to compute a transformation, they must also define the transformation between *all* points in an image.

Therefore, there is a need to control the regularity of this transformation either in an interpolation or an approximation context. This is crucial for medical imaging, since the geometrical transformation has a physiological significance (e.g. heart motion, optimally compressed breast etc.). In other words, if the transformation reflects only the true correspondences, the rest of the point (pixel/voxel) correspondences between images may be not be valid.

There are two possible approaches: to regularise the data before calculating the transformation, or to control the "smoothness" or the uncertainty in the localisation of all landmarks in an approximation scheme. In the case of field-based registration techniques (where correspondences are derived across the whole image), it may be more appropriate to regularise the correspondence data ("filter" the landmark set by rejecting outliers). For example, techniques exploiting correspondence "smoothing", Bayesian estimation of local error, Markov random fields, etc. can be used to correct or improve the local smoothness of correspondences resulting in a smoother transformation.

In the case of landmark-based methods, the previous approach to regularising the data may not be as appropriate since the transformation is calculated from these landmarks for the *entire* image. However the transformation can be smoothed by introducing a smoothness parameter in the calculation of the transformation. It is also possible to individually weight the "significance" of each landmark (e.g. by calculating its localisation uncertainty) [79]. In such a case, certain landmarks are given greater or lesser influence in the transformation calculation, hence reducing the error that may be introduced by landmarks that have a lower confidence. This concept is exploited in the next chapter where a two-stage registration algorithm for mammogram registration is presented.

Finally, there is a role for model-based registration in the context of regularisation. Shape constraints, deformable models with certain mechanical properties and smoothness conditions implied by assigning deformation characteristics to certain features (i.e. tissue classification in MRI) may also provide a measure of regularisation to a transformation.

The difficulty with such approaches is that the model must accurately reflect the characteristics of the registered images, otherwise the constraints may provide more error than benefit.

3.7.5 Optimisation

In many registration algorithms an optimisation step is included in order to improve the registration result and make the process more robust. In non-rigid registration applications, there are three main considerations related to optimisation:

Optimisation of correspondences: By including additional image-based information the calculated correspondences can be improved or "filtered" (e.g. the result of a search algorithm). For example, a collection of landmarks may be refined on the basis of feature parameters such as scale, orientation, shape, etc. to produce a more robust set of correspondences.

Optimisation of the transformation: This can be done either on the basis of iterative refinement or alternatively, control over the regularisation parameters. This is usually driven by a post-registration assessment of the quality of the transformation (either using a distance measure, covariance matrix, map of mutual information, etc.). In our application of non-rigid registration to temporal mammograms, we start with an interpolation approach when calculating a "boundary condition" registration, but switch to an approximation scheme when increasing the complexity of the deformation. Some optimisation examples include:

- Representing the transformation process within a framework that can be driven by alternative approaches. A good example of this would be to compute a fluid-flow displacement map but perform the transformation using radial basis functions driven by landmarks that are moved via the displacement [80].

- Optimisation in cases where the deformation (transformation) is based on "principal warps" or principle "modes" of deformation (e.g. in [76]). This applies equally well to radial basis functions (which may be treated as solution to an eigensystem problem) or to linear finite elements, where a deformation can be decomposed into a set of modes. By rejecting higher-order modes of deformation, the smoothness of the transformation can be controlled.

- Multi-scale optimisation: Many search schemes, both for point/region-based correspondence and displacement fields (such as those that result from optical flow) can be implemented in a multi-scale fashion. This optimisation issue addresses two considerations. Firstly, there is the issue of localising a landmark or feature correspondence to a particular region of the image using different scales of analysis. This may change the level of accuracy of correspondence and consequently the smoothness of the transformation (i.e. this generalises the localisation of a feature or displacement vector). In the case of optical flow, a high-scale evaluation of the displacement field will produce a much smoother approximation as it is effectively the result of summing a large number of smaller displacement vectors. The second optimisation issue that is addressed by multi-scale is computational efficiency. In image registration applications where large regions of the image (for example a fatty or involuted breast) do not contain much textural information, a multi-scale approach is likely to perform far better at searching for correspondences. Many multiscale techniques based around sampling windows, Gaussian (linear) scale-space, wavelets, etc. have been developed to address this problem. A multi-scale approach may also be useful for speeding the convergence of a search algorithm and to avoid local minima in the case of flow-type algorithms.

3.7.6 Interactivity

There are three classifications of registration algorithms with respect to interactivity:

Automatic: The user supplies the image data and possibly information about the image acquisition.
Semi-automatic: Interaction with the user is necessary either to initialise or to supervise (e.g. accept/reject landmarks) the registration.
Manual: The user defines the points or regions of correspondence.

Although automatic registration offers the greater challenge, very often in practice it is not the best solution for clinical related applications. In order to achieve automated performance, there is usually a trade-off between the complexity of the clinical application, the precision required and the degree of variability in registration that can be managed successfully. In addition, the clinician has to supervise and judge certain steps of the process in order to minimise possible errors. In mammograms, there is a large degree of variability in the imaging conditions, hence unless there are significant measures, a-priori knowledge and ultimately, heuristics, included in the registration framework, the potential to automate the process is usually limited. Clearly, by introducing heuristics, there may be a compromise in robustness. For all these reasons, it is essential to consider interactivity issues as well as robustness, speed and accuracy when designing the implementation of clinical registration systems and deciding upon the desired degree of automation.

In our opinion, there are three important considerations concerning medical image registration methods:

- The clinician must understand each step of the registration. To this end, the performance, the outcome and the clinical implications of the registration must be presented to the clinician (perhaps via visualisation or a graphical interface) in a way that is clearly understandable, and understood.

- The registration process must be capable of being interrupted and changed by the clinician at any stage. This means that although an algorithm may perform well in automatic mode, the clinician must be able to introduce his "expert" opinion into the

registration process to influence the final result, even if this degrades the result in some cases. This is the only way that the clinician will have faith that the process produces the outcome he/she desires.

- If the registration process fails, the clinician must be informed why, and at what stage (and due to what factors) the registration was not successful. This is very important since in many cases clinical data exhibit large variability and recording the reasons of registration failure is necessary to improve robustness. To this end, there must be measures incorporated into the registration process that evaluate whether or not the computed transformation is meaningful. An example of this might be a case where critical points (e.g. anatomical landmarks) cannot be detected due to occlusion.

With these considerations in mind, in this thesis, there has been an effort to develop registration algorithms that perform well automatically; but which also have a level at which the clinician can intervene.

3.8 Algorithms

3.8.1 Introduction

Having classified registration techniques, we examine several concepts and methods related to registration and we present some of the core registration algorithms in the literature. Most techniques require the calculation of correspondences between images. If the intensity distributions of the image pair are not substantially different, measures such as mutual information (discussed later in this section) can provide good matching results. In such cases, algorithms like optical flow and the demons algorithm (both discussed later in this section) may be used.

However, when different modalities are involved, or when the imaging conditions considerably vary (e.g. in X-ray mammography), the distributions of the image intensities may differ significantly. In such cases, photometrically invariant techniques are required

to segment similar structures in the image pair. It is necessary to include additional information either by approximating the intensity relation function of the images (from the joint histogram), or by establishing correspondences between image primitives using constraints or heuristics specific to the problem.

To this point, registration has been discussed in an abstract formulation. If we ignore the basis upon which the transformation is calculated (to a certain extent, this is the least critical component of the registration process) and focus instead on the derivation of correspondences between images, a series of commonly used matching approaches can be outlined. These include:

- Voxel similarity measures
- Fluid-based correspondence
- Optical flow
- Demons
- ICP, search-based approaches

There is an additional classification of "matching" that should not be discounted, namely the manual placement of landmarks. In cases where the transformation is simple (i.e. rigid/affine) and the visualisation/interaction of the medical data is appropriately flexible, it may be more appropriate for the clinician to define a small number of landmarks, than to automate correspondence detection.

3.8.2 Mutual information

Mutual information is a measure of similarity that has been used extensively in medical image analysis in order to calculate corresponding regions in images and to drive the registration process. It has been used as the basis of non-rigid multi-modal registration methods [81-83]. The mutual information (*MI*) of an image *I*, is defined as:

$$MI(I) = \sum_{i,j \in I} p_{i,j} \log \frac{p_{i,j}}{p_i p_j}$$

Where p_i, p_j are the probabilities of image intensities i and j respectively, and $p_{i,j}$ is the joint probability of i and j (the probability of a pixel transition from i to j in the image). The mutual information of two images I_1, I_2 is defined as:

$$MI(I_1, I_2) = \sum_{i \in I_1} \sum_{j \in I_2} p_{I_1 I_2}^{i,j} \log \frac{p_{I_1 I_2}^{i,j}}{p_{I_1}^{i} p_{I_2}^{j}}$$

Where $p_{I_1 I_2}^{i,j}$ is the joint probability density of intensity i in the first image and j in the second, and $p_{I_1}^{i}$, $p_{I_2}^{j}$ are the probabilities of the same pixels in their corresponding images. Viola [84, 85], was one of the first to develop an algorithm that calculated the transformation which leads to maximum mutual information between the source (model) and target image. Gilles [86], used the same concept including an optimisation scheme to avoid local minima of *MI* in the calculation of the best transformation. The exact transformation that maps the first image **u** (the model) onto the second image **v** (the image) should give rise to the largest mutual information. Mutual information then becomes an optimisation criterion, optimised with respect to **T:**

$$d/d\mathbf{T}[MI(\mathbf{T})] = d/d\mathbf{T}[\, H(\, \mathbf{v}(\mathbf{T}(X))\,)\,] - d/d\mathbf{T}[\, H(\, \mathbf{u}(X), \mathbf{v}(\mathbf{T}(X))\,)]$$

Where H is the entropy and *MI* the mutual information. The method uses a classic gradient-descent optimisation technique, and tries to find the transformation that gives the largest mutual information by taking small steps in the "direction" of the derivative of the criterion. This technique cannot be used directly in the non-rigid case since it is not possible to parameterise the induced transformation.

Mutual information can be used in a non-rigid registration framework to establish point correspondences (e.g. comparing the *MI* from patches along the two images). In [73], the

mutual information registration is presented as an iterative process, where each voxel u_n moves toward the gradient of the mutual information:

$$u_{n+1}(x) = G_\sigma \otimes (u_n + \alpha \nabla MI(u_n))$$

Where u_{n+1} is the calculated evolution of the point u_n and G_σ a Gaussian filter for smoothing. Comparisons of similarity measures for image registration can be found in [87, 88], while in [89] the correlation ratio is proposed as a new similarity measure.

3.8.3 Optical flow

The optical flow algorithm can be used to register temporal images providing that the intensity of each pixel or voxel $x=(x_1, x_2, x_3)$ remains constant in time [90]. In other words:

$$\frac{dI(x,t)}{dt} = 0$$

Using this constraint, the displacement $f(x)$ in the direction of the brightness gradient can be calculated as follows:

$$f(x) = -\frac{\partial I(x,t)/\partial t}{\left\|\nabla_x I(x,t)\right\|^2} \nabla_x I(x,t)$$

This equation is not sufficient to completely determine the transformation for all pixels, and many different suggestions have been offered to regularise the resulting vector field. Among these, Thirion [91], proposed to smooth the vector field with a Gaussian filter G_σ and to add a term in the denominator for numerical stability. With these considerations in mind, Guimond and Roche [73], showed that the optical flow method can lead to the following iterative registration framework:

$$u_{n+1}(x) = G_\sigma \otimes \left(u_n - \frac{\partial I(x,t)/\partial t}{\left\|\nabla_x I(x,t)\right\|^2 + [\partial I(x,t)/\partial t]^2} \nabla_x I(x,t) \right)$$

Where each voxel u_n is displaced to u_{n+1} using the previous recursive equation. This method is similar to the Demons algorithm introduced by Thirion [92]. As mentioned above, it is related to the optical flow algorithm by:

$$u_{n+1}(x) = G_\sigma \otimes (u_n - \frac{T(x) - S\,oh_n(x)}{\|\nabla T(x)\|^2 + [T(x) - S\,oh_n(x)]^2} \nabla T(x)$$

Where u_{n+1} is the calculated evolution of the point u_n and G_σ a Gaussian filter for smoothing, T is the target image and $S \circ h_n(x)$ the transformed source image in the n^{th} iteration. In [93], optical flow is driven by geometric rather than photometric criteria.

3.8.4 Fluid Registration

The original "viscous fluid algorithm" was a non-rigid registration technique introduced by Christensen et al. [94]. Several researchers have presented different implementations of the same idea [80, 95] either using multi-resolution optimisation, or by combining the method with thin-plate spline warps to obtain faster and more stable results. In the viscous fluid algorithm, the image S is deformed iteratively to approximate the target image T. At each step, the following partial differential equation (pde) has to be solved:

$$\mu \nabla^2 \mathrm{v} + (\lambda + \mu) \nabla (\nabla \mathrm{v}) + f(\mathrm{u}) = 0$$

Where μ, λ are the viscosity parameters, u the displacement vector in the source image and **v** the velocities corresponding to positions x-$\mathbf{u}(x)$ in the source image. The force f(u), that drives the deformation is given by:

$$f(\mathrm{u}) = -(\, S\,(\mathrm{x}\text{-}\mathrm{u}(\mathrm{x})) - T\,(\mathrm{x})\,)\,(\nabla S|_{\mathrm{x}\text{-}\mathrm{u}(\mathrm{x})})$$

At each iteration, the driving forces f(u) are calculated as the difference between the transformed source image (S) and the target image (T) multiplied by the gradient of the deformed source image (note the similarity with "demons"). After the forces are

calculated, the pde is solved at each step in order to calculate a field displacement u, for every image pixel. The basic problems with this method are: the computationally expensive implementation scheme that is required; and the "blurring" of the transformed image after several iterations, due to the regridding of the pixels. (Applying a transformation in an integer array such as a digital image can force the pixels outside the orthogonal pixel grid).

3.8.5 Combined registration-intensity correction

As mentioned in the introduction to this section, intensity-driven registration algorithms do not always give good results, especially in multi-modality registration problems. This is because the intensity distributions of the two modalities can be very different. However, Guimond and Roche [73], developed an iterative algorithm to register brain images that effectively reduces the problem to a mono-modality registration task. At each step, the intensity relationship between the images is approximated (assuming either mono-functional or bi-functional dependency) and the source image S is transformed into the "intensity corrected" S^*. After the intensity correction, they used a version of the demons algorithm to transform the geometry of S^* and better approximate the target image T:

$$u_{n+1}(x) = G_\sigma \otimes (u_n + \frac{S_n^* \circ h_n(x) - T(x)}{\left\|\nabla(S_n^* \circ h_n)(x)\right\|^2 + [S_n^* \circ h_n(x) - T(x)]^2} \nabla(S_n^* \circ h_n)(x)$$

Where u_{n+1} is the calculated evolution of the point u_n and G_σ a Gaussian filter for smoothing, T is the target image, S^* the intensity corrected source image and $S \circ h_n(x)$ the transformed source image in the n^{th} iteration. The algorithm converges to a solution, when the result is not significantly improved between two successive steps.

3.8.6 ICP-based algorithms

Many variants of the ICP (Iterative Closest Point) algorithm ([96, 97]) have been widely used for geometric alignment of 3D images. Registration methods based on the ICP algorithm give the advantage that they can include intensity, orientation, gradient as well as geometrical information in a minimisation scheme that does not require explicit calculation of derivatives. In general, the steps of the algorithm are:

- Selection of points to be matched between source and target image
- Establishing correspondence between points
- Weighting the calculated matches
- Rejecting pairs that may represent a matching error (e.g. due to occlusion or pathology present in one of the images).
- Minimise a cost function calculated from the matched pairs.

One basic weakness of these algorithms is that a good initialisation is often needed but this can sometimes be avoided with an initial rigid estimation of the transformation (e.g. based on mutual information).

In [98], both the intensity and geometry are aligned, deforming the image in 4D during iterations. The minimised energy is given by:

$$E(f,g) = \sum_{(M_j,N_j)\in Match_i} d((f(x_i), g(M_j)), N_j)^2 + \lambda \cdot \sum_{(M_j,N_j)\in Match_i} S(f,g,M_j)$$

Where f, g are the geometrical and intensity transformations respectively, x_i the spatial co-ordinates vector, S is a smoothness function, λ is the smoothness parameter, while M_j, N_j, and $Match_i$ are defined as follows*:

$$Match_i = \{(M_j, N_j), where\ N_j = CP_{4D}([f_{i-1}(x_j), g_{i-1}(M_j)]\}$$

* CP stands for closest point

In other words, for each iteration i, a set of closest points is defined between the source image (M) and the destination (N) and subsequently the functions by minimising the energy-cost function.

3.9 Registration Assessment and Validation

In cases where the registration is based on a rigid or rigid/affine approach, the validation is relatively straightforward. The performance of the registration is less dependent on the computation of the transformation (i.e. such techniques are simple and numerically robust) and more dependent on the choice of landmarks. In this case, clinician interaction will produce a registration framework that can be very robust. More automated approaches such as maximisation of mutual information, covariance matrices, etc. may also produce reliable registration assessments for simple transformations. A large-scale validation experiment of rigid registration has been reported by Fitzpatrick and West [99]. Various research groups registered the same image pair and the resulting transformation co-ordinates for three pre-defined points were used to assess the error in the results. The ground truth transformation calculated using fiducials.

For the non-rigid case, registration validation and assessment is substantially more difficult. In general, we are interested in the performance of the registration with respect to the following considerations:

- Robustness and stability: It is necessary to develop robust registration algorithms that can deal with data variability and do not allow "outliers" (e.g. a low confidence landmark pair) to change the stability of the approach. The numerical stability is important, since most of the algorithms calculate a significant number of complex tasks through the images. Generally, a large number of heuristics is not desired, since it reduces the robustness of the process and makes the whole process less intuitive.

- Precision and accuracy. This is particularly critical if an approximation method is used or a model is included to constrain the transformation. Does the registration

algorithm provide sufficient precision to be used for treatment planning, image guided surgery, etc.? If the precision fails, can the clinician be informed? Phantom studies and clinical evaluations that compare the results of registration with a "expert" opinion may be useful for evaluating this consideration.

- Generality of assumptions. Although this relates to stability and robustness, it must be clear that any assumptions or heuristics introduced into either the matching or transformation computation, must be valid for the entire range of clinical cases for which the algorithm will be used. If heuristics are critical to the success of the algorithm, can the validity of the heuristic be assessed for each case (and the clinician appropriately informed).

- Computational performance. The choice of registration algorithm may also be influenced by the computational performance of a particular approach. In safety-critical registration applications or cases where a diagnosis (that depends on the registration) is needed within a certain period of time, the quality of the registration may be offset by the time needed to utilise the ideal solution. A good example would be the general inappropriateness of using a PDE-based (partial differential equation) registration for intra-operative registration, compared with a simple search and radial basis functions.

- Clinical "ground truth". Most applications of registration result in the fusion of image data for the purpose of evaluating pathology, treatments or utilisation for surgical guidance. In all of these cases, the clinical "ground truth" is the pathology (or histopathology). The best way to evaluate registration algorithms is via localisations, image-guided biopsies, post-surgical assessment, etc. as this provides real evidence that the registration technique has been successful. Clearly this is not trivial, especially for non-rigid registration problems. In the case of multi-modal fusion (for example, breast mammography to MRI), this kind of localisation/biopsy study is critical due to the large extent of the deformation during registration.

3.10 Mammogram registration and matching: A challenging problem

3.10.1 Introduction

The most important mammographic component in terms of disease diagnosis, is the fibroglandular tissue. Mammographic images are hard to interpret and, since the region of interest consists entirely of soft tissue that is compressed during acquisition, the mammographic appearance is not always consistent. The radiographer tries to optimally compress the breast so that all the glandular tissue elements are separated during acquisition. For this reason, in the registration work presented in this thesis, the glandular tissue is considered to be the most important image feature, and the internal matching is based on the detection of glandular structures (work presented in chapter 4). Other structures, like curvilinear structures, vary significantly between successive scans and are not considered to be reliable features for registration.

Figure 11 shows a typical L-R pair of cranio-caudal mammograms. The most prominent characteristic is the high intensity architecture consisting of the glandular tissue and Cooper's ligaments. The high intensity architecture appears as a bright network against a dark background, which consists of fatty tissue. The appearance of x-ray mammograms varies according to the age of the woman. A normal, young woman's mammogram is usually bright due to the high-attenuation characteristics of calcium that is prominent in the milk ducts of the breast. As a result, the detection of abnormal structures in younger women's mammograms is today considered very hard.

Mammograms obtained from postmenopausal women vary in appearance according to the degree of involution that has occurred. From the standpoint of physiology, involution is defined as gradual degradation of milk-bearing tissue to fat (which is trans-radiant). As far as the mammographic appearance of the breast is concerned, radiolucent areas appear in the place of high intensity regions in the mammogram. Such images that contain both high intensity regions and radiolucent areas can be confusing when the radiologist has to consider the presence of an abnormality.

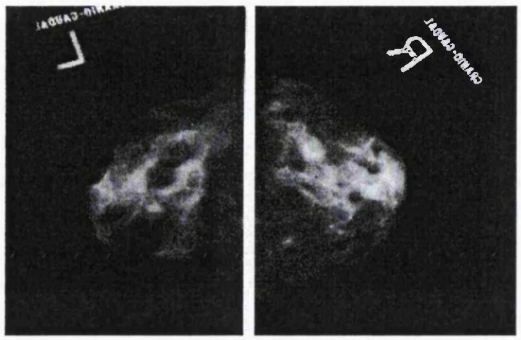


Figure 11: A bilateral mammogram pair

Temporal mammogram registration is a mono-modal and, to date, intra-subject registration problem. It is a complex problem since the breast is a particularly dynamic organ and its composition varies according to a large number of parameters (diet, stress, hormones, age, disease, surgery etc.). In addition, the compression of the breast before each mammogram is taken can vary between successive acquisitions. Although, in absolute value, the difference in compression does not vary significantly (typically around 0.5cm) the resulting differences in the geometry of the projected tissue can be significant. As has been shown in differential compression experiments [4], a 0.5cm difference in compression can change the mammographic appearance disproportionately (curvilinear structures can "vanish", while tumours can move a long way).

Figure 12 illustrates the most common temporal changes that can make registration very difficult:

- Changes in compression and patient positioning result in changes in the projected area of dense tissue.

- Changes in imaging conditions make equivalent parts of the breast look brighter or darker in one of the two images.

- Breast tissue can change dramatically over the years (involution, HRT, weight gain/loss). In addition, the presence of the pathology or changes due to surgery/treatment can significantly alter the appearance of the breast.

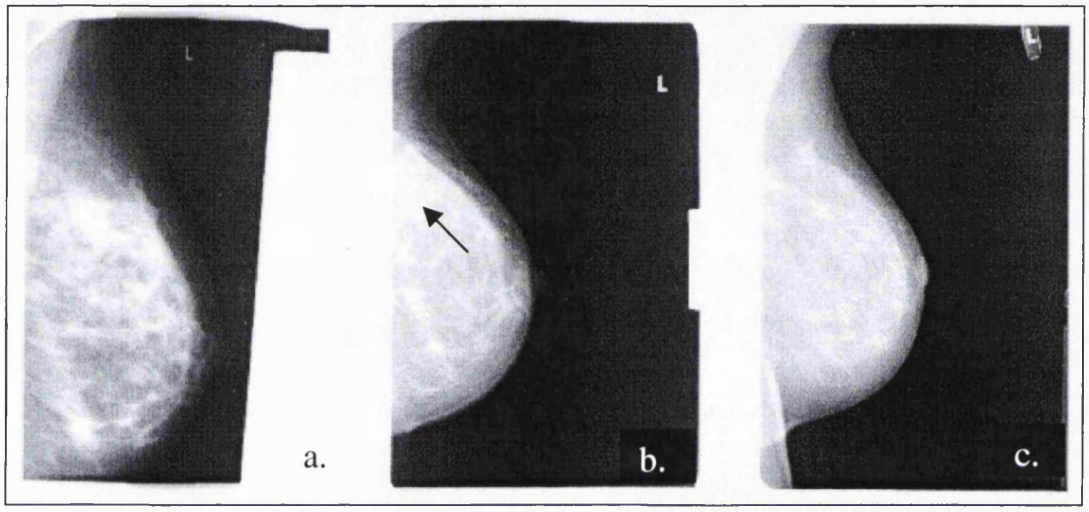


Figure 12: Temporal images of the same patient, a) is the earliest and c) the most recent. The combinations of changes in the structure inside the breast, in compression and in the imaging conditions make the registration problem a non-trivial process. In addition, after cancer was diagnosed (in the upper outer quadrant (UOQ) first image from the left) it was excised. The scar tissue (seen as the bright region where the arrow points) can cause confusion when reading the mammogram.

In this section, the reasons that make temporal mammogram registration a hard problem in medical image analysis are explained in detail. These consist of the changes in breast compression, imaging conditions, new growths and changes due to HRT.

3.10.2 Differences in breast positioning and compression

The first reason for dissimilarities in the appearance of a X-ray mammogram between sessions is the fact that the breast is positioned differently in the X-ray machine. This

corresponds to a rigid displacement but is straightforward to compensate for (by translating and rotating the image).

The breast compression often varies between two successive acquisitions since it is only recorded and controlled weakly. The reason is that in order to produce a diagnostically useful image, the radiographer tries to compress the breast so that the tissue structures are separated as much as possible. As a consequence, the tissue structures can move more (with additional compression) and the resulting projected areas of tissue can be very different. In other words, even a small change in compression between a pair of temporal mammograms, can lead to a very significant change in the locations of dense ("bright") regions.

An increase in compression leads to a global expansion of the breast tissue that corresponds to a relatively smooth flow [4, 100]. However, the overall smoothness is disrupted by tissues that are significantly more dense (e.g. tumours) and by local tissue interactions.

Highnam and Brady analysed in detail the different aspects of breast compression in their differential compression mammography studies [4]. Figure 13 shows an example, where even a small change in compression significantly changes the appearance of the mammogram between two consecutive (differential mammography) acquisitions. Several of the considerations presented in their work, are related to the difficulties in registering mammograms. Differences in the properties of the various tissue types in the breast (e.g. stiffness) lead to a non-uniform displacement of the breast structures under compression. Cysts are soft and mobile, while fibroadenomas are soft and may be deformed under pressure. Finally, because of their higher density cancers are barely deformed and do not roll under breast compression. These ideas were exploited in [4], where the breast tissue was imaged successively by slightly changing the compression. It was shown that under additional compression cancers moved significantly less than the surrounding fibroglandular structures, making their detection easier than detecting them from a single

mammogram. Such sequences (differential compression) can be used to validate the intensity correction (after registration) method suggested in chapter 5.

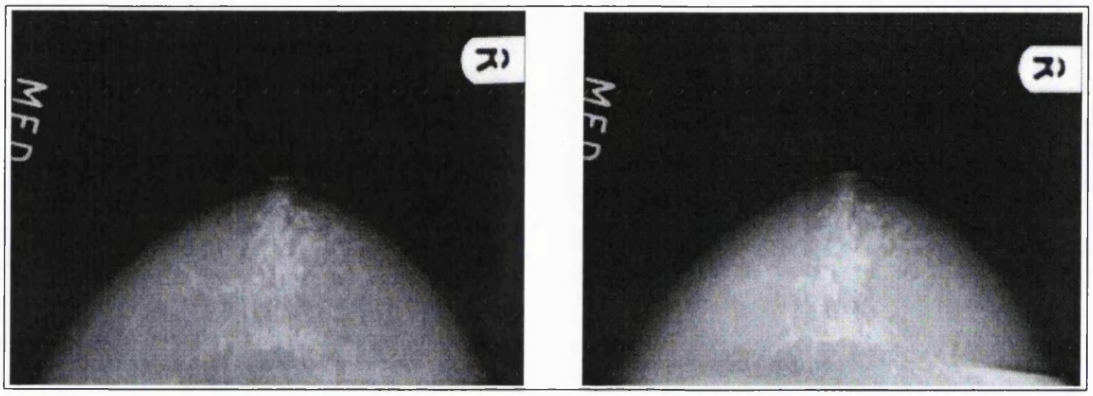


Figure 13: Slight increase in the compression (~0.5cm) and change in imaging parameters (e.g. time of exposure) significantly change the appearance of the mammogram, even thought the patient is kept immobilised between the two successive acquisitions [4].

3.10.3 Differences in imaging conditions: The need for intensity normalisation as well as geometric alignment

Because of the relatively weak control over the image acquisition process, it is difficult to eliminate variability in image characteristics, such as contrast and brightness. Such differences in imaging conditions lead to a non-rigid transformation between the intensities of temporal pairs of mammograms. For this reason, it becomes even more difficult to register temporal mammograms, consequently to assess temporal changes by comparison. In addition, the existence of intrinsic degrading factors (scattered and extra-focal radiation) reduce the diagnostic value of each mammogram.

The h_{int} representation of interesting tissue (discussed in chapter 5 and in Appendix C), introduced by Highnam and Brady [4], is a quantitative measure in which each pixel represents the thickness of 'interesting' (non-fat) tissue between the X-ray source and the image. For that reason, it can be used to normalise mammograms since it provides

anatomical information about the breast. Using the h_{int} representation, Highnam and Brady, modelled and removed the effect of scattered, extra-focal radiation as well as image glare [101] thus creating a general framework for mammogram normalisation.

The h_{int} representation is particularly useful in the case of HRT sequences, where the objective is to assess local changes in the composition of the breast tissue, as mammogram matching and registration do not factor out variations in the pair due to imaging conditions. In HRT temporal mammography, both intensities and the geometry of the mammograms needs to be aligned. Since the imaging conditions often vary temporally, even if a pair of images is accurately aligned, differences in the intensity of corresponding points inevitably remain and only by using the h_{int} representation one can achieve a quantifiable comparison of a mammogram with previous ones. For this reason, in chapter 5 of the thesis the two sources of information are combined in order to develop a robust technique for assessing HRT changes.

3.10.4 New growths

Breast cancer is a very common disease, and one of the reasons for its high mortality rate is the difficulty in early diagnosis due to the complexity and variability of mammograms. Any new growth present in the most recent mammogram, either malignant (cancer) or benign (cyst, fibroadenomas), reduces the architectural similarity with previous ones. This is useful for establishing diagnoses via comparison of successive mammograms; but it renders the image registration problem more difficult. Figure 14 illustrates this consideration.

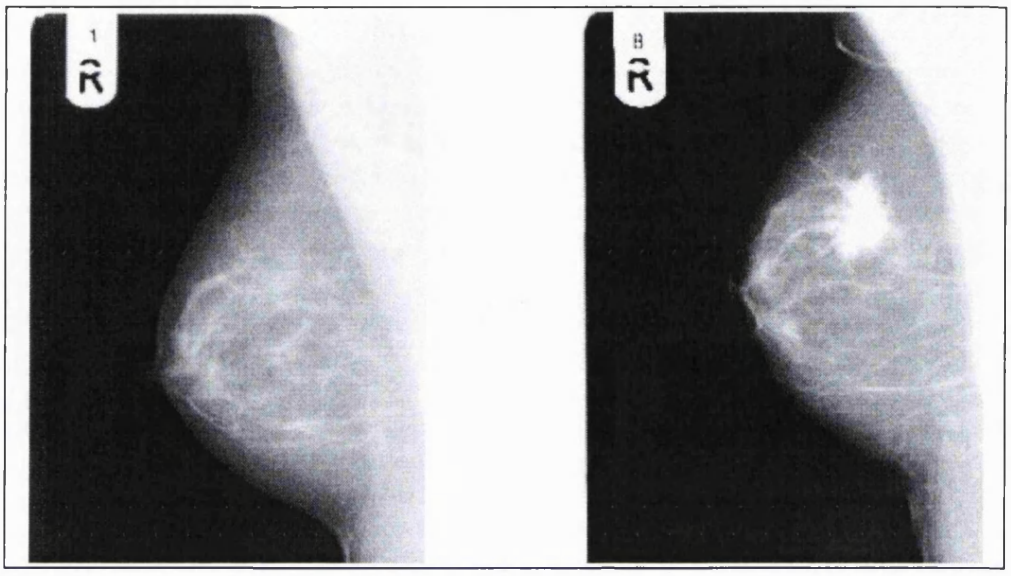


Figure 14: A cancer has been diagnosed in the most recent mammogram of a patient (right image). The registration problem becomes more difficult (decreased similarity between the images), but also more clinically significant since the comparison can help to establish diagnosis and retrospectively evaluate if there were signs of the cancer in the previous mammogram (left image).

3.10.5 Hormonal-Related changes

In younger women, the breast tissue structure varies daily due to the hormonal changes that occur during each menstrual cycle. In post-menopausal women, hormonal deficiencies cause the gradual transformation of fibroglandular tissue into fatty tissue, and again the mammographic appearance of the breast can change dramatically. The reverse process (tissue regeneration) can occur to women that are treated with Hormonal Replacement Therapy (HRT), making the interpretation of mammograms more difficult and the establishment of temporal correspondences harder. The work in mammogram registration and tissue quantification for HRT mammogram sequences is presented in chapter 5.

3.11 Previous work on mammogram matching and registration

3.11.1 Introduction

As mentioned in the introduction, registering a mammogram with a previous one is made more difficult by possibly different compressions of the breast at the two time instances, and the likely differences in the imaging conditions. The effects of these differences, which from now will be described as temporal changes, result in a non-rigid transform (geometric) in the image plane due to difference in compression, and a non-rigid transformation (photometric) in the intensities of the images. In this section, previous work on mammogram matching and work on bilateral mammogram registration is discussed.

3.11.2 Mammogram matching

3.11.2.1 Siew-Li Kok-Wiles work

Kok-Wiles's work on establishing correspondences between mammograms, is based on the observation that if no pathology is present, the parenchymal patterns should not differ greatly between the mammograms [100, 102]. Consequently, differences or movements observed in the dense tissue regions should be due mainly to involution processes and to new growths as well as to differences in x-ray imaging conditions, breast positioning and compression.

The key innovation in Kok-Wiles' work is to represent the breast as a set of salient regions, then match the salient region structures across images. There are three aspects to the representation:

1. **Regions** of increased brightness in the mammogram are the basis of the representation. They are chosen because breast changes, particularly interval cancers, appear as regions. Representing and matching the regions means that a cancer appears as a region that cannot be matched successfully, or as a region whose representation

(area, contrast, texture, etc) have changed significantly (e.g. corresponding to significant growth of a tumour);

2. **Salient** regions are those which appear *a priori* to be good candidates for matching. There are, quite simply, too many regions that can be extracted from a mammogram, the search-space is very large. For this reason, Brady and Kok-Wiles devised a scheme based on the contour of the candidate region as well as its average contrast to the background to filter out "obviously" irrelevant regions.

3. **Nested** regions are represented explicitly, forming a tree-like structure of regions. This is a useful representation of the topological structure of the bright regions in the breast. Brady and SLKW argued that it is largely preserved over time and for left-right matching.

Novak's compression study [103], was re-interpreted to yield several constraints that were used to model the non-rigid temporal transformations:

- Topological constraint: It is assumed that the nested structure of salient regions is very similar in both images.
- Smooth motion constraint: The transformation of neighbouring regions is similar. This assumption comes from the fact that relative movements of tissues during breast-compression are not large or sudden.
- Order and monotonicity of geometry constraint: The order relationships between geometrical properties of regions (e.g. area, contour length) are preserved in the temporal pair.

Using these constraints regions are matched between temporal or bilateral mammograms [100]. A potential pathology in one of the images can be identified as a salient region with no good match. The main problems of this method is the large number of salient regions (iso-contours that fulfil the saliency constraints) and heuristics necessary for the matching.

3.11.2.2 Vujovic and Brzakovic's work

Provided that further (than those in Kok-Wiles' work) assumptions are made, it is possible to try and match points belonging to vessel intersections or, more generally, to the curvilinear structures of the breast. This is the case in N. Vujovic and D. Brzakovic work [104, 105], where they select intersections of image structures such as ducts and blood vessels as potential control points which they subsequently match between the mammograms.

Although their algorithm is successful at detecting such points even in predominantly radio-opaque mammograms, there are robustness issues concerning the use of this method for matching temporal mammograms, since it has been shown that the curvilinear patterns change significantly as a result of: the difference in compression exerted on the breast; and involution [4]. Although sometimes vessel structure is a good sign of architecture preservation, it is susceptible to differences in compression and imaging conditions. For that reason, such a method for acquiring control points in mammogram pairs would only be acceptable if the pattern of the curvilinear structures remains constant.

Such a method would not be appropriate if temporal changes like involution or new growths have occurred. On the other hand, Kok-Wiles' algorithm, although it is highly dependant on heuristics [102], is more generic than the N. Vujovic one, since it matches iso-intensity contours corresponding to bright regions (glandular tissue, masses). Additionally, Kok-Wiles algorithm is independent of curvi-linear structures, whose appearance can change dramatically under different compression in subsequent (or bilateral) scans.

3.11.3 Landmark based Registration techniques in mammography

3.11.3.1 Introduction

The utilisation of thin-plate spline interpolation as a way to recover deformations in medical images was introduced by Bookstein [76]. In its original version, the method did not allow uncertainty in the location of the landmarks, and for that reason it was modified by [77]. It has been used in registration of heart images [106] as well as bilateral breast registration ([107], [108]). This method is efficient at recovering local deformations; but special care is needed in the selection of the landmarks. The calculated interpolating function $f(x, y)$ for the vertical or the horizontal direction is smooth and deforms the image in such a way that the bending energy is minimised, while the landmarks are matched. The general form of this function is given by:

$$f(x,y) = a_1 + a_2 \cdot x + a_3 \cdot y + \sum_{i=1}^{n} w_i \cdot U\left(\left|P_i - (x,y)\right|\right)$$

Where P_i are the n landmarks on the boundary of the first mammogram, α_1, α_2, α_3 and w_i are the n+3 coefficients calculated by the method. The resulting interpolating function is calculated by minimising the bending energy of the transformation, given by the next equation:

$$I_f = \iint_{R^2} \left(\left(\frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \cdot \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 \right) dxdy$$

Thin-plate spline interpolation and approximation are used in our mammogram registration algorithm, presented in the next chapter. This method is efficient at recovering local deformations; but special care is needed in the selection of the landmarks. In particular, landmarks in the inside of the breast are difficult to extract reliably, and errors in their localisation can degrade the accuracy of the overall transformation.

After aligning the two images by using the calculated transformation, difference images can be produced in order to look for regions of large intensity differences. These regions can either be new growths (e.g. expansion of a cancer), changes due to involution, or due to local inaccuracy in registration. These issues are explored in the next chapter.

3.11.3.2 Previous work

Karssemeijer and te Brake [107] and Bowyer and Sallam [108], have presented work on registration of bilateral mammograms using thin-plate splines for the automatic detection of abnormalities through visual examination of the difference image after registration. Breast asymmetry is the main characteristic that can be assessed using the method. However, both techniques use points spaced evenly along the breast boundary, along the chest wall, and include points from the pectoral muscle line. In addition, the method for the automatic detection of corresponding points presented in [28] cannot be robust, since in most cases it is impossible to establish point-to-point correspondence in a bilateral pair of mammograms. We contend that special care is needed in order to decide which points should be used for mammogram registration and this issue is addressed in the next chapter.

Bowyer and Sallam [108], used transformation constraints to reduce the search-space of potential matches. This was done in a two-stage interpolation. The first stage calculates the interpolating function that aligns the two images based on corresponding points along the breast boundaries of the two images. In that way, most of the global (rigid), as well as some local deformations near the breast boundaries, is recovered. Moreover, the initial warping function is used in the next step to define the regions in the second image, where the algorithm will locate the best matches for the selected features of the first image. Ultimately, combined corresponding features from inside the breast and along its boundary are used to calculate the final interpolating function.

It is not clear from their published work if the second stage of interpolation improves the initial one. They neither present images showing the matches nor images showing the improvement in the interpolation at the second stage. In addition, as will be discussed in the next chapter the large number of boundary points is unnecessary and the utilisation of landmarks on the pectoral muscle "line" can be problematic.

On the other hand, Karssemeijer and te Brake's work [107], does not take into consideration that specific correspondences are required to match the breast boundary in the two images (for example the nipple, as the glandular structures converge to the nipple [8]). In addition, they align the approximated pectoral muscle lines without including internal landmarks, something that is proved to be unreliable in the next chapter.

3.11.4 Conclusion on previous work

In the previous sections we summarised relevant work on mammogram matching and registration. We believe the work in this thesis addresses the following problems:

- Previous work on mammogram matching and registration was mainly concerned with *bilateral* mammogram pairs and did not address the specific issues for temporal registration. In this thesis we analyse in detail the problems related to *temporal changes* on mammogram sequences.

- Previous work on registration based on the boundary did not offer a method for reliably establishing correspondences in the breast boundary. As is discussed in the next chapter the linear expansion of curvature used in [108] does not realistically address the problem. In the next chapter, a reliable method for aligning the boundary in mammogram pairs is presented. Moreover, the multi-scale segmentation and subsequent matching of internal structures suggested can improve the registration result.

- Lastly, previous work showed a very limited number of results and applications. In this thesis the proposed two-stage algorithm is used for temporal and bilateral mammogram registration including HRT sequences. Results are also shown in the Gallery (Appendix E). In addition, the possible applications (e.g. interval cancers or assessment of local density changes) of such methods are discussed in detail in the next chapter.

3.12 Conclusion- Summary

In this chapter, the basic issues involved in a registration problem have been presented. The distinction of the necessary steps, categories and methods is not always clear but it is important to have an understanding of the general registration problems before the mammography-specific techniques are presented in the next chapter.

In addition, the most commonly used algorithms for mono-modal or multi-modal medical image registration were discussed. Although, in many cases (like mammography) it is not possible just to use directly a registration method since the problem is quite complex, it is important to have a good understanding of registration techniques, since some of them (or a combination of them) could inspire the reader to further improve the mammogram registration method presented in the next chapter.

Lastly, the reasons that make mammogram registration a hard problem were discussed. Not only are the projection images not always consistent, but changes in the breast (due to body weight, water content, involution, disease, surgery, long term effect of exogenous hormones etc.) can make the architectural similarity vary a lot and the image alignment a hard task. Previous work in mammogram registration does not offer a reliable framework for aligning mammograms and maximising the similarity between the images. In the next chapter we explain in detail the contribution of this thesis in mammogram registration, taking into consideration the limitations and the problems related to image acquisition.

Chapter 4: Proposed Registration framework for mammography

4.1 Introduction to the proposed method for mammogram registration

Clinical motivation

As was discussed in the first two chapters of the thesis, the reliable diagnosis of abnormalities from a single mammogram is very difficult, and so it is increasingly the case that pairs of mammograms are compared. These might be left and right mammograms taken at the same session. Also, in the reading centres, if earlier mammograms are available, the radiologist will compare them to the most recent one(s). This process can be especially important for:

- Women at high risk of developing breast cancer (e.g. women with family history of breast cancer or genetic susceptibility-BRCA-1*), usually have more frequent mammograms taken in order to detect a malignancy at as early a stage as possible. Previous ("normal") mammograms are used as a baseline for comparison with recent ones.

- Post-menopausal women often decide themselves, or are advised by their GP, to undergo Hormone Replacement Therapy (HRT). However, despite the many beneficial effects of HRT, has been reported that long term use leads to an increased risk for developing breast cancer. As is discussed in the next chapter, there is a suggestion, based on clinical experience, that localised increases in tissue density could be an early indication of breast cancer for women using HRT. For this reason, it is important to be able to register and compare HRT mammogram sequences, aiming again at a more effective comparison for early diagnosis of cancer. The next chapter deals with temporal mammography on HRT patients and addresses the crucial issues of accurate tissue quantification and temporal/differential analysis.

* BRCA-1 is the "breast cancer gene". Several of its mutations have been found to cause breast cancer.

- Retrospective studies aim to analyse temporal data in order to assess the accuracy and effectiveness of diagnosis in hospitals/screening centres. Such studies aim to define the rate of missed cancers and interval cancers, as well as to further educate clinicians in the vitally important task of early diagnosis. Since diagnosis through comparison is, in many cases, crucial (especially when new growths are subtle), alignment of mammogram sequences can assist retrospective studies and training.

The aim is to develop temporal registration tools that can **assist the clinician to compare mammograms more efficiently**. As was discussed in the previous chapter, the comparison of a mammogram with a previous one is made more difficult by differences in compression of the breast at the two time instances and the likely differences in imaging conditions. The impact of these differences, which from now on will be described as *temporal changes*, consist of a non-rigid transform in the image plane due to difference in compression (even a small difference in compression can lead to a significant and uneven displacement of the breast structures), and a non-rigid transformation of image intensity due to differences in the acquisition conditions. In addition, normal changes (e.g. weight gain or loss, involution) as well as affects of Hormone Replacement Therapy or chemotherapy agents such as Tamoxifen can introduce more differences in the architecture of a temporal pair. The intensities of a mammogram pair can be normalised using the h_{int} representation of interesting tissue (Appendix C). In this chapter, we discuss in detail our method for the *geometrical alignment* of mammogram sequences.

The domain and the aims of the proposed work

The work in this chapter aims at the development of a robust framework for the geometrical alignment of temporal (or bilateral) mammograms. Due to the non-rigid motion of the different breast tissues during compression, we need to consider a method that computes a pixel to pixel non-rigid transformation between the two input mammograms. Ideally, after geometrical alignment each voxel in the two input images should correspond to the same column of tissue of the compressed breast; this is however

subject to changes in the breast (e.g. tissue involution, cancer). The aims of this chapter can be therefore summarised in the next points:

- We aim to develop a method that will exploit the structural similarity of any given temporal mammogram pair, in order to calculate a geometrical transform for image alignment. The algorithm we aim to develop should be useful for clinicians in comparing any mammogram pair. More specifically:

  - The registration result should never be unreasonable because this would undermine the clinician's confidence.

  - The alignment result should be close to the clinician's perceptual judgement. For example, a point to point mapping based on the registration should be close to the expert's assessment. This is further discussed in section 4.5.5.

- In order to achieve the above, we need to determine suitable classes of landmarks which can anchor the geometric registration.

- We also aim to determine whether or not there are additional useful constraints that can be used in the registration process. For example, the pectoral muscle has been used in mammogram registration but without questioning its effect in the registration result.

- Explore the possibility of applying the same method to bilateral mammogram registration and discuss its possible clinical applications.

The working hypothesis

The main hypothesis for the proposed work in this chapter is that it is possible to determine a non-rigid geometrical transformation in such a way that the similarity of

normal tissue samples will be increased in the registered images; and it is possible to find patient related landmarks automatically for each patient.

Previous work on bilateral mammogram registration has shown promising results using the breast outline for mammogram registration. However, none of the previous methods offers a robust and intuitive method that enables the selection of consistent boundary landmarks for the *automatic alignment* of the breast boundary. This is discussed further in section 4.2.3.1. The breast outline offers useful constraint on the geometric transformation but that is insufficient since breast tissues move non-rigidly under compression, and in such a way that reflects their mechanical properties.

The mammographic appearance of the breast varies significantly between acquisitions due to compression, variable imaging parameters and the actual changes in breast tissue. Nevertheless, we believe that most of the times it is possible to define a set of internal correspondences that can be used to refine the registration result and improve it. This hypothesis is inspired from the perceived "architectural similarity" that is most often the case in temporal mammograms.

Description of the selected methodology (why this method was chosen)

The problem of mammogram registration was outlined in the previous chapter. It was concluded that none of the existing methods could robustly align both the boundary and the significant internal structures in mammogram sequences. The method proposed in this chapter was developed in order to overcome the main problems in temporal mammogram registration which are due to a combination of a non-rigid tissue motion due to different compression between acquisitions, differences in the imaging parameters and the temporal changes in tissue composition and structure of the breast. In section 3.10, we presented in detail the reasons that make temporal mammogram registration a difficult task. Figure 12, illustrates most of these changes in an example of temporal mammogram sequence. Since the intensity distribution and the structural morphology can significantly differ in temporal sequences, we need a *photometrically invariant* method that can

exploit the *variable* similarity of temporal pairs of mammograms. For this reason, altough the registration techniques presented in section 3.8 are very powerful, they can't be used in any direct way for mammogram registration.

Next in this chapter a three-stage method is proposed to register temporal sequences of mammograms:

Initially the images are aligned based on the boundary. This is done by developing an algorithm that automatically detects 3 points with characteristic curvature in the outline of both mammograms. A thin plate spline interpolation is used to calculate the image transformation that aligns the boundaries of the two mammograms. The method for the robust calculation of consistent landmarks using curvature analysis in the breast outline to align the images using thin-plate spline interpolation is described in section 4.2.

Using a wavelet-analysis segmentation algorithm we define internal regions of dense tissue that have good spatial characteristics in each mammogram. This is discussed in section 4.3. The boundary transformation together with scale and area information of the segmented regions is used to match internal structures and refine the registration. In this second optimisation stage, a regularised approximation scheme is used to account for possible inaccuracies in the selection of the internal landmarks, especially because the center of mass of each matched region pair is used to calculate the image transformation. The approximation step of the registration is discussed in section 4.4.

All the software in the thesis was developed in the C++ programming language and the time required to register two 300 micron digitised mammograms is 45 seconds on a PII 500MHz machine. This time is required for calculating the geometrical transformation and for warping the second input image in the co-ordinate system of the first one.

Since there is a large variability both in the information content of each mammogram (e.g. highly dense breast or involuted breast) and in the architectural similarity of mammogram pairs, internal matching is not always possible. A good example is the case

of fully involuted breasts, where the only internal structures are vessels which cannot be used as landmarks since their mammographic appearance changes dramatically under different compression [4]. Figure 15, is a flow chart of our registration algorithm. Though the technique could be fully automated, the acceptance or not of the internal landmarks should in practice be decided or confirmed by the user. If the suggested internal landmarks do not meet the clinician's satisfaction (e.g. possibly in involuted breast pairs) the boundary-based registration is the final result, otherwise an approximation scheme (including internal and boundary landmarks) is employed in order to better approximate the deformation necessary to align the mammograms.

In summary, in this chapter we first present a mammogram registration technique based only on boundary points to undo most of the effects of different compressions is presented. Then, it is considered how to improve registration by using internal landmarks before presenting results, validation and a discussion concerning the possible applications of registration. A gallery of results is presented in Appendix E, showing the application of the proposed registration method in bilateral and temporal mammogram pairs.

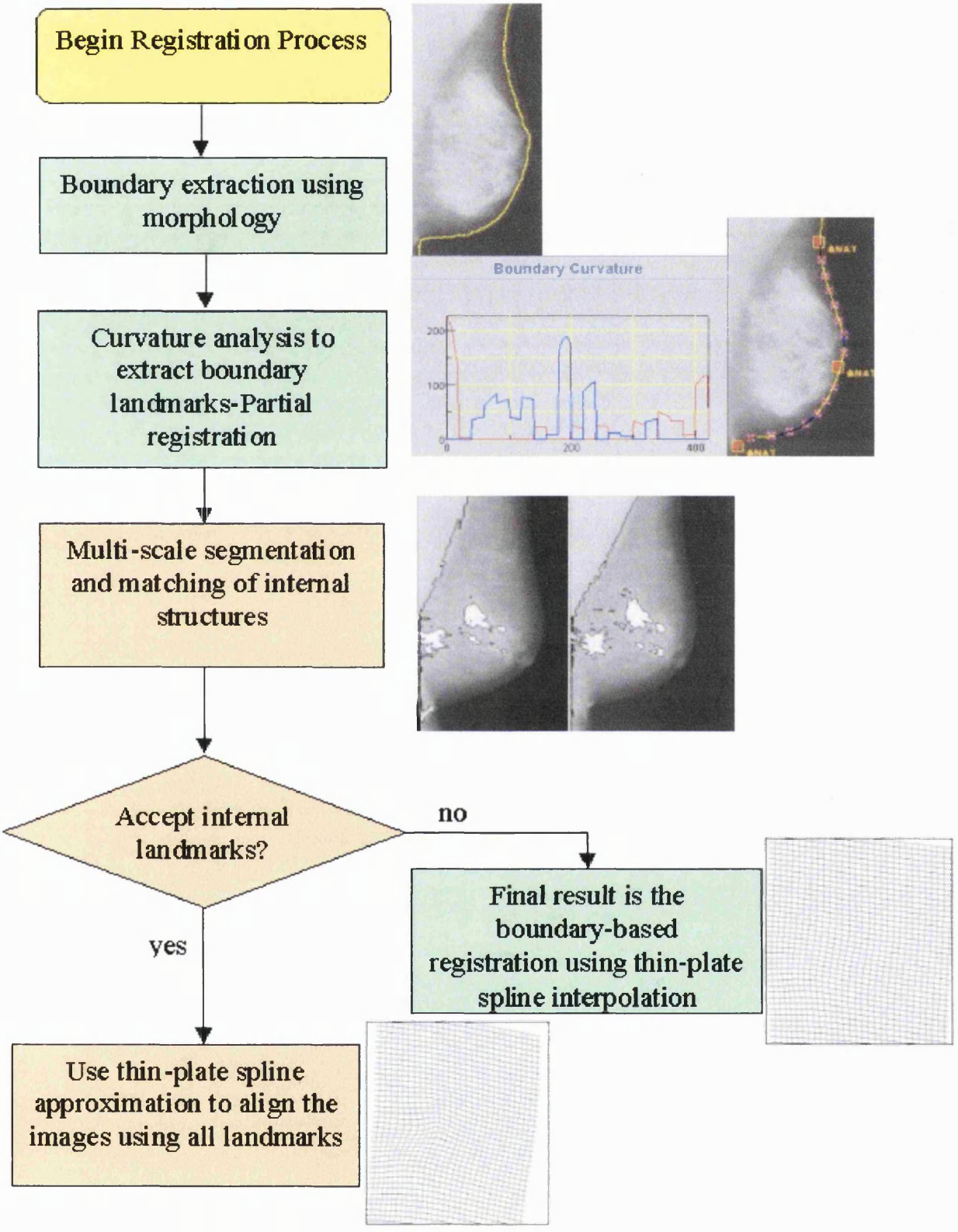


Figure 15: The basic steps of our breast registration algorithm

4.2 Partial Registration using the boundary

4.2.1 Introduction

The breast boundary is the most useful feature of the mammogram in terms of temporal consistency. It provides information about the difference in compression between two acquisitions and enables the calculation of landmarks that allow the approximation of the transformation that relates the geometry of a temporal pair of mammograms. As will be discussed in the next section, the registration using the boundary is not sufficient as internal structures move to different extents under different compressions, because of differences in shape and tissue density. However, accurate detection of the breast outline and calculation of temporally invariant geometrical landmarks is a key first step for mammogram registration. The steps that comprise the boundary registration method are:

1. Boundary outline detection (section 4.2.2)
2. Curvature analysis of the outline(s) and detection of consistent landmarks (section 4.2.3)
3. Anatomical significance of detected boundary landmarks
4. Thin-plate spline interpolation to align the boundaries (section 4.2.4)

These steps are explained in detail before addressing the need for internal landmarks for improved registration.

4.2.2 Breast outline detection

To generate a segmentation, the image is thresholded in the first "valley" of the intensity histogram as shown in Figure 16(a), resulting in a binary image I(b) (Figure 16(b)). Subsequently, an 8-connected component outline is obtained using mathematical morphology (closing ($\bullet$) followed by dilation ($\oplus$) and then subtraction). The relation between the binary I(b) and the outline I(c) is given by the equation:

$$I(c) = (I(b) \bullet M_{disc}) \oplus M_{square} - (I(b) \bullet M_{disc})$$

where M stands for a mask (the radius of M_{disk} is 5 pixels and the size of M_{square} 3 pixels). An example is shown in Figure 4.2(c).

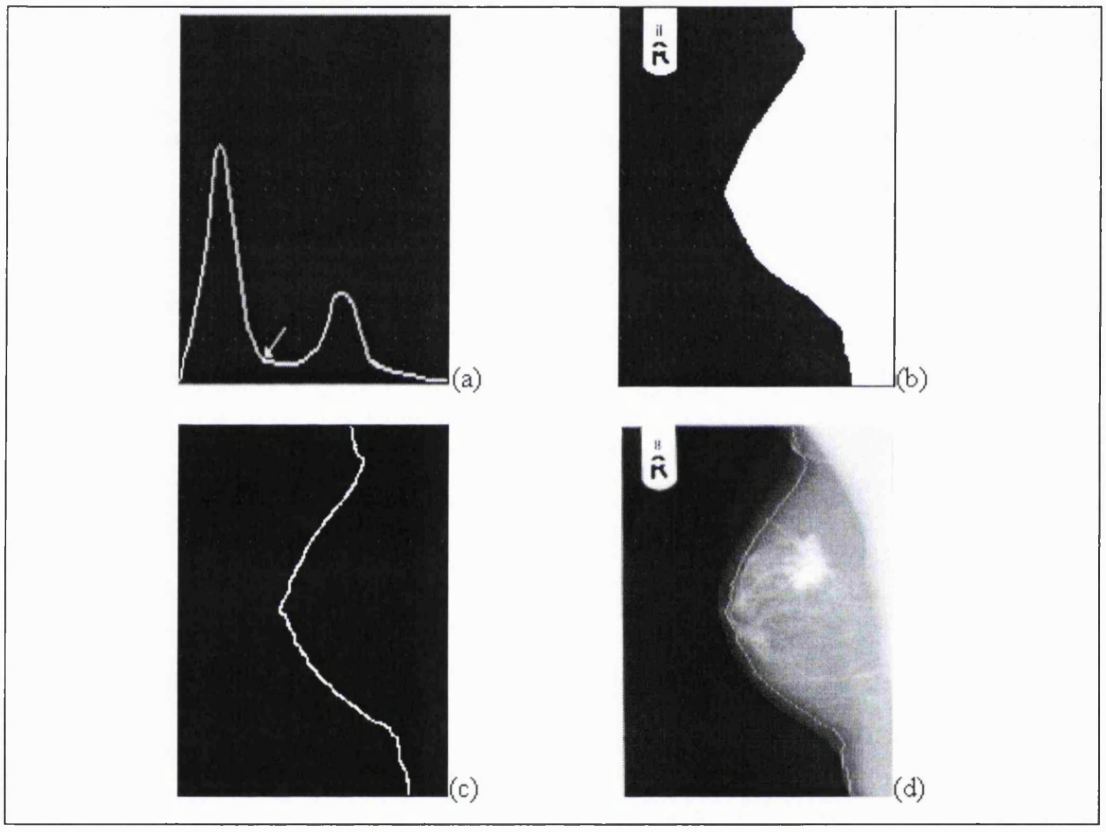


Figure 16: (a): Approximating the breast edge using the histogram of intensities, (b) The binary image after theresholding, (c): The boundary is extracted as the largest connected component in the image after using morphology, (d): The outline superimposed in the original image.

This method yields an approximation to the boundary, it is very fast and the resulting outline is an 8-connected component curve (due to morphology). This serves our purpose of easily tracking the points along the boundary since the breast outline can have an irregular shape (e.g. very often 2 points in the *x* axis can correspond to 1 in the *y*).

However, the segmented curve is "jagged" and this poses a problem for the robust detection of boundary landmarks, as is discussed in the next section.

4.2.3 Curvature analysis on the breast outline-consistent landmarks

4.2.3.1 Defining consistent boundary landmarks using curvature

In order to be able to register mammogram pairs one needs to be able to establish correspondences between the breast outline of each mammogram, since the "beginning" and "end" of each outline highly depend on the segmentation result and the image acquisition (e.g. in some medio-lateral mammograms a larger part of the rib is visible than in others). The aim is to "translate" the geometrical consistencies in mammograms into an automatic algorithm for the detection of *consistent* boundary landmarks. Bowyer and Sallam [108], and Karssemeijer and te Brake [107] have presented work on the registration of bilateral mammograms based on the boundary. However, neither method calculates consistent points along the boundary (in [108] a curvature linear expansion algorithm is used to match boundary points, while in [107] the whole segmented outline is sampled without establishing any boundary correspondences). Due to the non-rigid expansion of the breast during compression, the assumption of linear expansion in the curvature of the breast outline is not valid in many cases (e.g. in mammogram pairs where the outline is almost in the same place in the upper part, from the nipple to the axilla, while is expanded in the lower one). Nevertheless, it has been observed that a small number of "pseudo-invariant" points can be detected.

Figure 17 demonstrates this idea, showing three points of characteristic geometry in the "ideal" breast outlines. In the cranio-caudal case (CC), the points 1 and 3 can be assumed (as in [108]) to be near in the chest wall (and thus invariant) and are approximated by the first and last points of the breast-outline respectively. Point 2, is the maximum curvature point (negative curvature by convention). The medio-lateral mammograms (ML), represent the most difficult case. However, we have found that the three points shown in Figure 17 are consistent for almost all ML mammograms (e.g. there is only a few cases, especially in very large breasts, where the axilla is not visible). Geometrically, these

points can be described as two maxima of positive (by convention) curvature (points 1 and 3) and 1 point of maximum negative curvature (point 2).

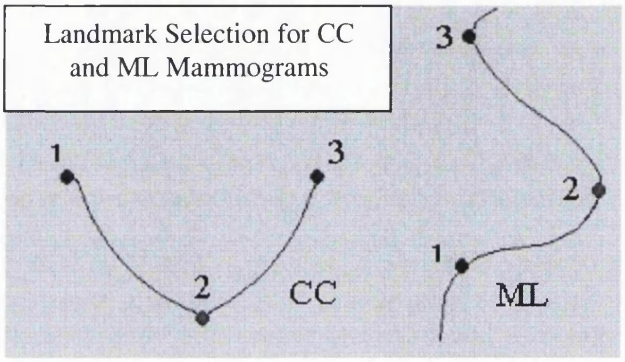


Figure 17: Consistent landmarks in the CC and ML "idealised" outlines.

In order to build a robust detection algorithm for the three points discussed above (in the remainder of this section we will refer to the ML case only, as it is more general than the CC) **we need to calculate the curvature profile of the breast outline**. However, there are some intrinsic limitations:

- The segmentation method used approximates the outline using a global threshold and it is not always smooth. This results in a large number of local maxima (and minima) of curvature along the breast outline.

- The calculation of curvature involves the estimation of second order derivatives which lead to "noisy" curvature profiles and makes the robust detection of the suggested points hard.

To overcome such problems, Asada and Brady [109], have suggested a Gaussian multi-scale analysis of features in 2D curves. A similar concept is proposed:

1. Design an algorithm that automatically detects the three suggested points (Figure 17) based on an "idealised" model of the breast outline. This algorithm is based on the separation of positive and negative curvature as is described later on this section.

2. Define an optimum sampling rate (S_{opt}) of points along the segmented breast-outline and run a spline to approximate the breast boundary. Different sampling rates preserve different amount of detail at a trade-off with overall smoothness. This optimum would depend on the pixel dimension (300 microns in the images we used). The aim of this step is to approximate the idealised outline shape (shown in Figure 17) and be able to robustly detect the suggested 3 points.

These two steps are now described in detail:

Step 1: An algorithm for the automatic detection of high curvature points

Once the breast outline has been calculated, it is sampled and approximated with a spline $\boldsymbol{r}(t)= (\boldsymbol{x}(t), \boldsymbol{y}(t))$ which is a continuous function computed for regularly sampled points (we use every three) of the breast outline. As is discussed later in this section, the spline resolution ("density" of interpolated points) does not affect the detection of the anatomical landmarks. However, the sampling of the breast outline is *crucial* for the detection of anatomical landmarks. To build the algorithm, it is assumed that the outline is "ideal" (as in Figure 17). The optimum sampling S_{opt} is defined in the next step.

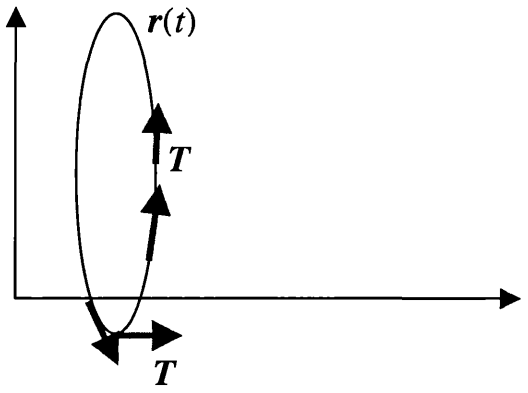


Figure 18: An example of a parametric curve $r(t)$. The rate of change of the unit tangent vector T is the curvature.

Figure 18 is an example of a parametric curve $\boldsymbol{r}(t)$. The curvature at each point is defined as the rate of change of T with respect to arc length S:

$$k = \left|\frac{dT}{ds}\right| = \left|\frac{\frac{dT}{dt}}{\frac{ds}{dt}}\right|$$

where $\frac{ds}{dt} = |r'(t)|$ is the speed, and T the unit tangent vector. For example if $\boldsymbol{r}(t)=\cos(t)\cdot\underline{x} + \sin(t)\cdot\underline{y}$, then $ds/dt= (\sin^2(t) +\cos^2(t))^{1/2}=1$ and $T=(\sin(t)\cdot\underline{x} - \cos(t)\cdot\underline{y})/1$, and the curvature can be calculated as: $k=dT/dt= -\cos(t)\cdot\underline{x} - \sin(t)\cdot\underline{y}$.

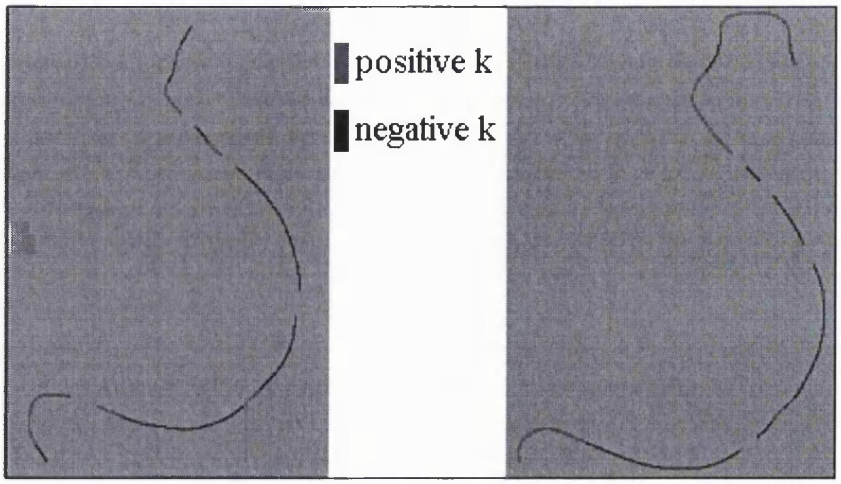
This way, the curvilinear parameter t describes the breast outline and we are able to calculate geometrical measures along the boundary. By using the spline approximation of the boundary, where each point is $(x(t), y(t))$, we calculate the curvature in each point t of the curve, according to the well known equation:

$$k(t) = \frac{x_{tt} \cdot y_t - x_t \cdot y_{tt}}{x_t^2 + y_t^2}$$

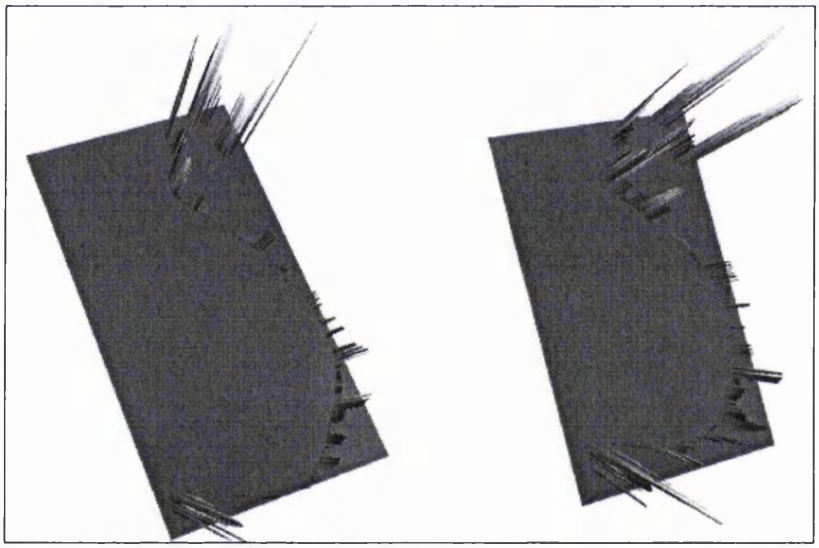
Where x_t, x_{tt} and y_t, y_{tt} are the derivatives of $x(t)$, $y(t)$ with respect to t.

The key to the maximum curvature detection algorithm is separating the positive and negative curvature segments of the outline. Figure 19(a), illustrates the positive and negative curvature segments in a temporal pair, while Figure 19(b) shows the positive curvature in 3D.

Figure 20, shows how the maximum curvature detection algorithm works. The curvature signal is unfolded like a 1-d signal and separated into the complementary positive and negative segments. Point 2 (referring to the idealised model in Figure 17), is calculated as the global maximum of "negative" curvature and after that we are able to look for 2 "positive" curvature maxima (points 1 and 3 in Figure 17) in each direction of point 2, using a 1-d maximum-detection algorithm.



(a)



(b)

Figure 19: (a): The positive and negative curvature segments in a temporal pair, (b): The positive curvature and the maxima (rib, axilla) in 3D

All the points of the "unfolded" curvature signal are assigned a value of 255. We start from the top of the graph (which is the highest value of curvature after rescaling the calculated values for all the breast edge) and each time a maximum value is detected a zero value is assigned in a zone neighbouring to the peak. That way, we avoid detecting neighbouring peaks as maxima, which ensures additional confidence (in addition to the fact that the peaks can't be in the same direction from the pre-calculated negative

curvature maximum) that points 1 and 3 are correctly calculated. The width of the zone W_z has been empirically determined to be:

$$W_z = \frac{1}{2 \cdot (spres)} + 15$$

We provide the following definitions:

- *spres*: is the resolution of the spline, which is calculated for every three sampled points of the breast outline. It defines how many points will be interpolated between each 3 sampled boundary points.

- S_{opt}: is the boundary sampling that preserves the details needed to calculate the 3 suggested anatomical points. In other words is the sampling that resembles the "ideal" curve of Figure 17. Around that optimum value the detection of the 3 suggested points does not change significantly.

By calculating W_z, the detection of curvature maxima does not depend on the "density" of the interpolated points. However, as mentioned previously, it *does* depend on the selected sampling resolution of points in the breast outline. In order to robustly estimate the three suggested points, the optimum sampling needs to be estimated, as is described in the next step.

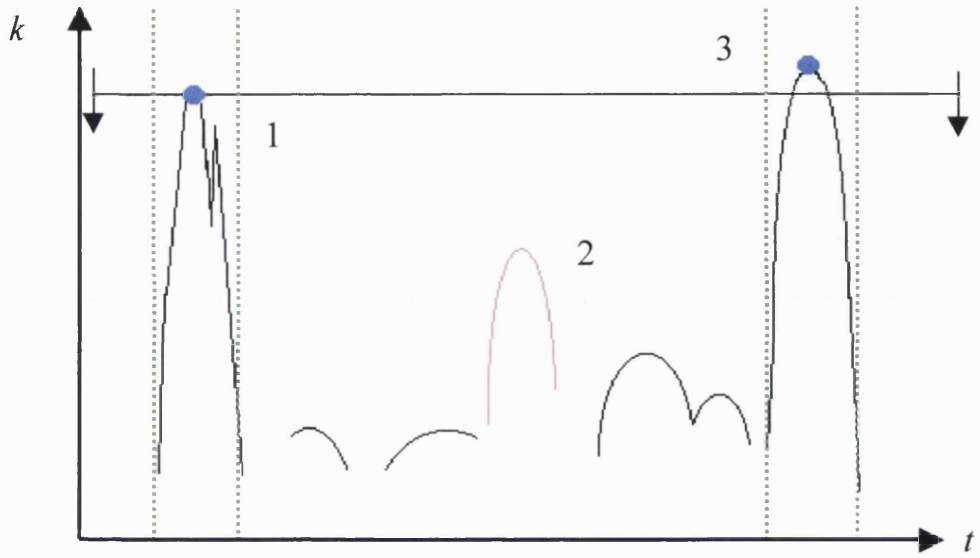


Figure 20: The maximum curvature detection algorithm. Point 2, is detected as a global maximum of negative curvature while to detect the two positive curvature peaks (points 1 and 3) we search from the top to the bottom and each time a maximum is detected (represented by the thick dots in the figure) the signal is "erased" in a zone around it so that neighbouring peaks are not detected.

Step 2: Defining the optimum sampling for the spline representation of the outline

There is a need to define the optimum sampling S_{opt}, in order to robustly detect the 3 points of characteristic curvature. This sampling depends on the size of the pixel (the images used in this thesis are 300 microns but this method can be extended to any resolution) and it can be defined as the sampling with which the maximum curvature algorithm described above converges to a steady solution. In other words, the sampling rate is increased until the maximum curvature points detected do not change significantly.

This is shown in Figure 21, where the detected points change significantly from sampling rate 15 (a sample is taken every 15 points along the 8-connected outline) in Figure 21(C), to 25 in Figure 21 (D), 35 in Figure 21(E) and finally 40 in Figure 21 (F). The last figure approaches the "ideal" outline shown in Figure 17, and the detected points reflect the

"perceived" points of maximum curvature (positive and negative) of the original mammogram boundary shown in Figure 21 (A).

In addition, the unfolded curvature profile approaches the "ideal" one (shown in Figure 20) as well. This is illustrated in Figure 22, where the "noisy" curvature profile in Figure 22 (a) (that corresponds to Figure 21(D)), is smoothed in Figure 22(b) and in Figure 22 (c) has approached the "ideal" case as mentioned before.

Finally, to define the optimum sampling S_{opt}, we plot the location of the detected curvature maxima along the outline for a range of samplings (from 10 to 50 with a step of 2), as shown in Figure 23. This is effectively a multi-scale extension of the maximum curvature algorithm described in the previous step. As is shown in Figure 23, after the sampling value of 40 the detection of the maximum curvature points converges. Using this value, 50 pairs of mammograms (of pixel size 300microns) were registered based on the boundary. The boundary registration technique (using thin-plate splines) is described in section 4.2.4. The success of this method can be attributed to the fact, that most of the time, the detected points of maximum curvature have an anatomical significance, as is discussed in the next section.

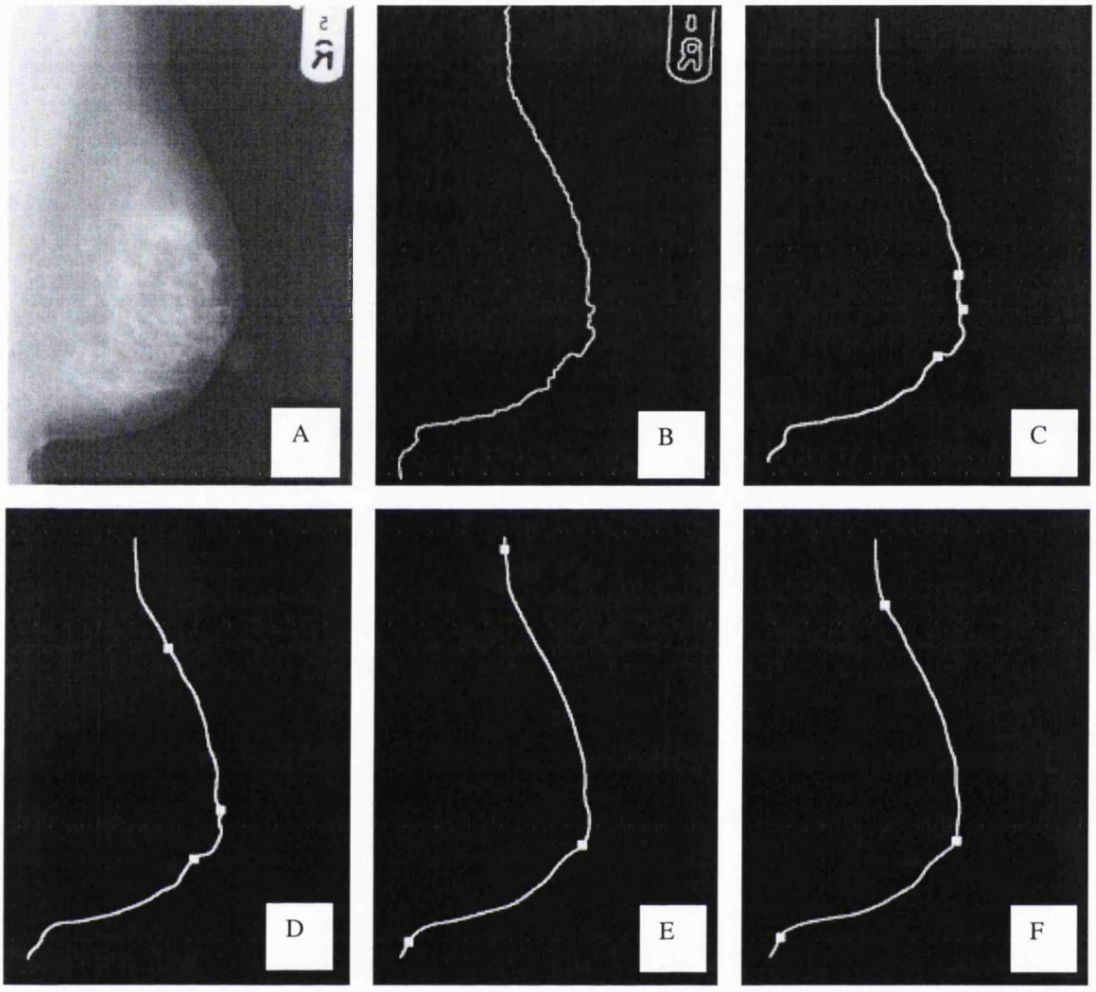
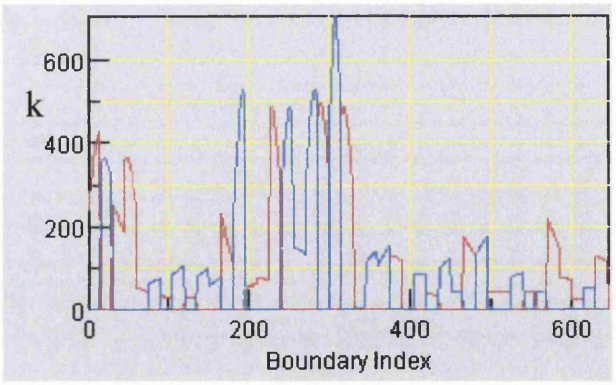
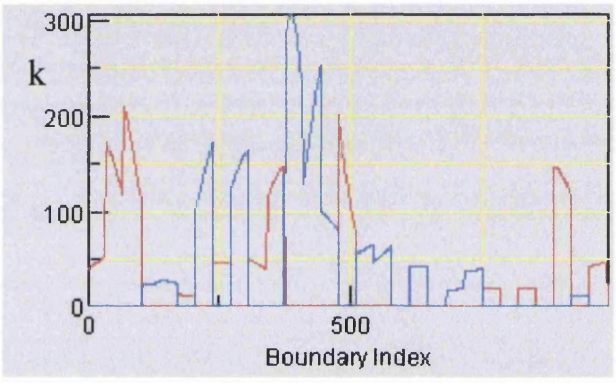


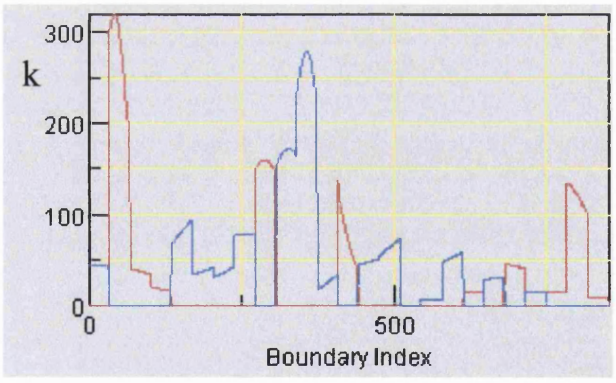
Figure 21: for the mammogram shown in (A): The original mammogram, (B): The breast outline. Note the "jagged" appearance due to the global threshold used, (C): The maximum curvature detected points using a sampling rate of 15 points along the boundary, (D): The same for sampling rate 25, (E): The same for 35 and (F): for 40. Note that the detected points change significantly.



(a)



(b)



(c)

Figure 22: The "unfolded" curvature profile for (a): Sampling rate of 25, corresponding to Figure 21(D), (b):Sampling rate of 35 corresponding to Figure 21 (E) and (c): Sampling rate of 40 corresponding to Figure 21 (F).

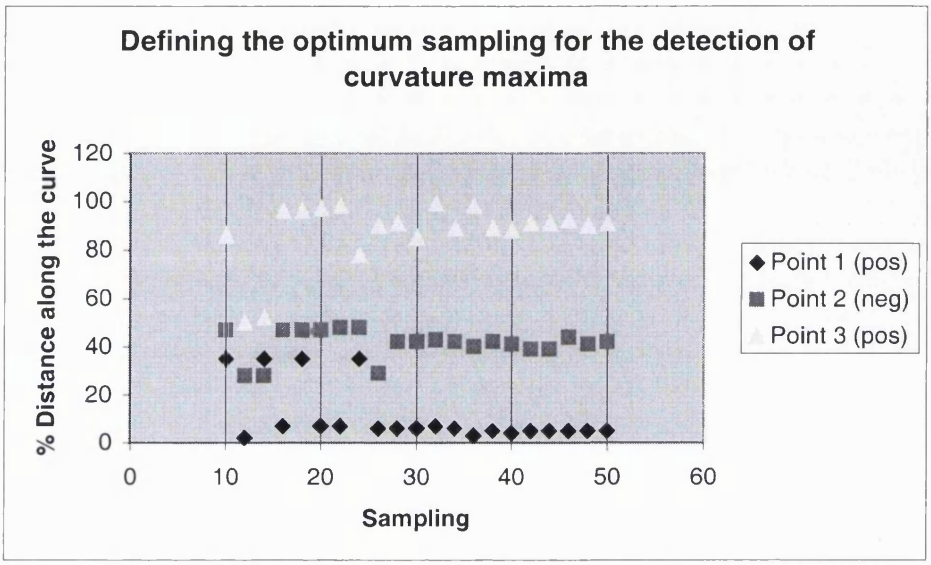


Figure 23: Defining the optimum sampling S_{opt} at the point where the detection of curvature maxima converges to a steady solution (at sampling rate of 40 in this case). Before and after this sampling, the point detection does not change significantly.

4.2.3.2 Anatomical significance of boundary landmarks

As mentioned previously, the boundary registration technique is based on the robust detection of 3 points of characteristic curvature (Figure 17). This makes the boundary alignment more robust as consistent boundary points are calculated for a mammogram pair, instead of attempting to match the whole segmented breast outline (as is done in [107]). The reason that this method works well is that in general these landmarks have an anatomical significance. As is shown in Figure 24, the detected landmarks very often correspond to the anatomical location of the rib (point 1 in Figure 17), the nipple (point 2), and the axilla (point 3). However, this anatomical correspondence is *not* a requirement for the boundary registration to work. A good example is the case where the nipple is not visible but there is always a global maximum of negative curvature.

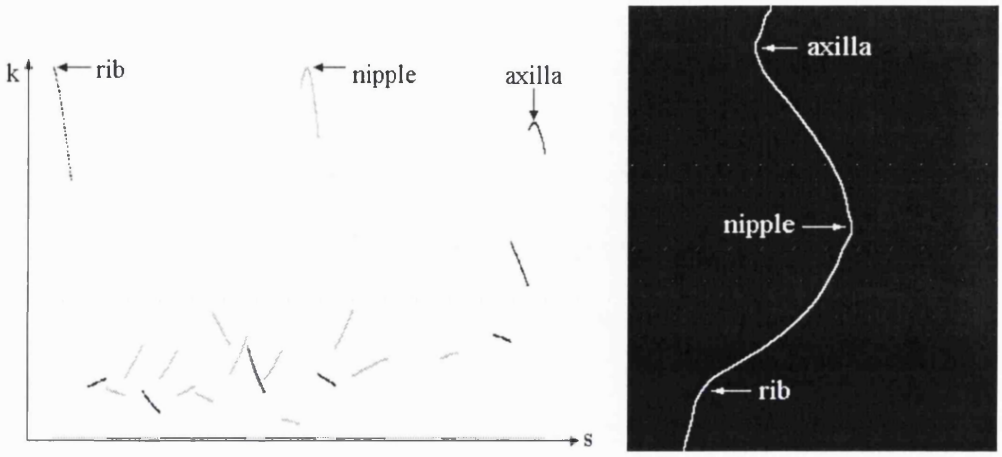


Figure 24: The selected points in Figure 17, very often have an anatomical significance. In the case shown, the curvature peaks along the breast boundary k (*s*) correspond to the rib, nipple and axilla points.

We believe that this algorithm could be the basis of a CAD system for the automatic detection of the nipple, axilla and rib points. To assess the anatomical significance of the detected points in the 50 mammogram pairs that were registered, we evaluated the correspondence between the detected landmarks and the location of the rib, nipple and axilla. Table 2, summarises results from 50 mammograms (medio-lateral and cranio-caudal). We assess the success in detection of these points individually (e.g. it is possible in a mammogram to correctly detect the axilla and rib points but not the nipple).

In our implementation, the user can refine the location of the boundary landmarks by shifting the calculated points along the breast-edge. As has been explained in chapter 3, effort making the registration process completely "blind" can have an effect in the robustness and adaptability of the method. In this context, the user might prefer to always register the images after making sure the 3 boundary landmarks correspond exactly to the corresponding anatomical points (although this would not necessary for the boundary registration to work!). The most important correspondence is the nipple, since the glandular tissue converges to it and, the method presented could be a starting point of a robust nipple detection algorithm. Although this is not exploited in this thesis, as is discussed in chapter 6, it could improve the registration method suggested.

Anatomical correspondence of detected landmarks in breast boundary	Correspondence rate	Additional detection by changing the sampling	No correspondence	Reasons for no correspondence
Axilla	75%	15%	10%	Axilla occluded, or is very close to the radiological tag
Nipple	65%	15%	20%	Nipple is not visible in the breast outline
Rib	85%	10%	5%	Rib point is occluded

Table 2: Assessment of the anatomical significance of the three detected points using the maximum curvature detection algorithm.

Boundary registration seems to vastly improve the correspondence of a mammogram pair. However, the registration can be improved if internal correspondences are included in the registration process and this is discussed later in the chapter. The next section explains the partial registration of a mammogram pair using the three detected points.

4.2.4 Partial registration from the breast boundary

4.2.4.1 Overview

Radial basis functions have previously been used to align bilateral mammogram images [107, 108]. The method described in this thesis reliably calculates boundary landmarks, which most often have an anatomical significance. Sampling between these three points (that from now on will be *only* referred as rib, nipple and axilla) and more specifically, between the axilla-nipple and nipple rib segments, any temporal or bilateral mammogram pair can be aligned based on the boundary.

For temporal mammogram registration we have observed that a good initial alignment can be achieved using at least five points along the breast boundary. However, for greater accuracy in aligning the boundaries, seven points uniformly sampled between the "axilla" and "nipple" boundary landmarks, and another seven between the "nipple" and the "rib"

landmarks (total of 17 points) are used. Using these points, the images are registered using thin-plate spline interpolation [76].

Radial basis functions (RBFs) are used for the elastic image deformation in this registration scheme. RBFs are used in two contexts – firstly, for aligning *only* the boundary. Later, when internal landmarks are included, information about the spatial characteristics of the deformation points is used to implement a more sophisticated regularisation that is based on an approximation method. We start with the interpolation case as it is more general to defining a RBF.

4.2.4.2 Interpolative Radial Basis Functions

In RBF interpolation, a set of n landmarks (p_i, q_i) *is* used to define a transformation function $u:R^2 \rightarrow R^2$, where $p_i=(x_{i1}, y_{i1})$ are the landmarks in the first image, $q_i=(x_{i2}, y_{i2})$ are the landmarks in the second, and the interpolated transformation function $u(x)$ must satisfy the interpolation constraint:

$$u(p_i) = q_i, i = 1,\ldots,n$$

To ensure smoothness of u we ensure also that it minimises a functional (regulariser) $J(f)$ for R^2.

$$J_2^2(f) = \iint_{R^2} \left\{ \left(\frac{\partial^2 f}{\partial x^2} \right)^2 + \left(\frac{\partial^2 f}{\partial y^2} \right)^2 + 2 \cdot \left(\frac{\partial^2 f}{\partial x . \partial y} \right)^2 \right\} . dxdy$$

This functional is also known as the bending energy of the deformation. In this case we are solving the biharmonic equation $\Delta^2 f=0$, whereby local deformation conforms to a function $f(x,y)$ for minimisation of the bending energy. For interpolation, each co-ordinate is calculated independently (i.e. the interpolation $u_k : R^d \rightarrow R$ is solved for each co-ordinate $k=1\ldots d$, in this case $d=2$, with the corresponding constraints $u_k(p_i) = q_{i,k}$). In the general RBF approach, the interpolation function $u(x)$ consists of two parts [74-76]:

$$u(x) = \phi_s(x) + R_s(x)$$

where $\phi_s(x)$ is a sum of polynomials up to degree p and $R_s(x)$ consists of a sum of RBFs:

$$\phi_s(x) = \sum_{j=1}^{M} \beta_j \phi_j(x)$$

$$R_s(x) = \sum_{i=1}^{N} \alpha_i R(\|x - p_i\|)$$

The $\phi_j(x)$ are a separable basis of M functions for all polynomials up to degree p. In the case of 2D images (e.g. mammograms), it represents the *rigid* part of the deformation. $R(x)$ is a function that depends only on the distance $x \geq 0$ from the origin, and describes the local deformation (interpolation) of points around the RBF landmarks. Moreover, $\| x - p_i \| = \| r \|$ denotes the Euclidean distance from x to p_i and, α_i, β_j are coefficients, with the RBFs $R(\| x - p_i \|)$ centred on the N landmarks p_i. Combining the last two equations and using the deformation constraint:

$$\sum_{i=1}^{n} \alpha_i \phi_i(p_i) = 0, \quad j=1...M$$

yields a system of linear equations for the stacked vectors of weights α and β:

$$\begin{pmatrix} K & P \\ P^T & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} q_k \\ 0 \end{pmatrix}$$

where: α and β are the vectors containing the weights α_i and β_i, K is the n x n matrix $K_{ij} = R(\| p_i - p_j \|)$ and P is the N x M matrix $P_{ij} = \phi_j(p_i)$. The $q_k = (q_{k,1},...,q_{k,n})^T$ are the set of target landmarks.

A wide variety of radial basis functions has been proposed in the literature for elastic deformation of images (e.g. in [108] for breast registration). These include thin-plate splines (R_{TPS}) [76], multiquadrics (R_M) or inverse multi-quadrics (R_{IM}) [76] and Gaussian RBFs (R_G) [74]:

$$R_{TPS}(r) = \begin{cases} r^{4-d} \ln r^2, 4-d \in 2N \\ r^{4-d}, otherwise \end{cases}$$

$$R_M(r) = (r^2 + c^2)^{\mu}, \mu \in R_+$$

$$R_{IM}(r) = (r^2 + c^2)^{-\mu}, \mu \in R_+$$

$$R_G(r) = e^{-r^2/2\sigma^2}$$

In our work to date, we have used the thin-plate spline radial basis function R_{TPS} since it is a stable method to recover deformations (including local deformations due to breast tissue motion). By weakening the interpolation constraint, the smoothness of the transformation can be controlled and the uncertainty in the localisation of landmarks can be taken into consideration. This concept is exploited in the second stage of the registration process. Bookstein [76], introduced the utilisation of thin-plate spline interpolation as a way of recovering deformations in medical images. The calculated interpolating function *f(x, y)* for the vertical or the horizontal direction is smooth and deforms the image in such a way that the bending energy is minimised, while the landmarks are matched.

Once the interpolating function has been calculated, "warped' images are produced by forcing every point *(x, y)* in a mammogram to take the intensity value of the point where the interpolating function maps the *(x, y)* point of the previous mammogram. Bilinear interpolation can be used to calculate intensities outside the pixel grid (as in [108]). After image warping, difference (subtraction) images can be generated and used to search for regions of large intensity differences. These regions can be either new growths (e.g. a cancer) changes due to involution, or they can be due to local inaccuracy in registration. Registration can be improved if we are able to automatically detect landmarks inside the breast tissue.

Extensive experimentation has convinced us that registration based only on the breast edge does not model with sufficient accuracy the relative motions of important tissue regions under different levels of compression. The boundary transformation accounts for the global differences between the images, correcting scaling (due to differences in compression or breast-size) translation and rotation due to breast positioning and orientation of the glandular structures (since the nipple is included as a landmark). However, as described in [100], different breast compressions (which are almost inevitable), tend to make denser structures (e.g. tumours) move far more than less dense tissues (like cysts or fibroglandular structures) resulting in disruption to an otherwise

generally smooth motion field. Indeed, such motion discontinuities are the basis for Highnam and Brady's differential compression technique, which aims to estimate tissue elasticity [4]. The feature detection described in the next section aims at an intensity-invariant segmentation of "important" regions that can be matched and thus allow us to better approximate internal tissue deformations by including internal landmarks is the registration process.

4.3 Multi-scale landmark selection for improved registration

4.3.1 Introduction

In this section, the need to establish correspondences between regions in temporal mammograms for robust and more accurate registration is emphasised. Based on automatically detected boundary landmarks, we partially register, then subsequently analyse the mammogram pair using a non-linear wavelet scale-space to isolate significant regions of interest. It is demonstrated that a usually small, but significant number of internal correspondences greatly improves registration and better approximates the complex internal tissue deformation due mainly to differences in compression.

To complete the registration process, corresponding internal landmarks are detected in the registration pair and, along with various matching conditions (discussed in the next section), are used to compensate for the complex internal deformation of the breast. We cannot rely on a field (intensity-based) registration technique as the deformation may be quite large, the projection angle can vary between temporal scans and the morphology of the breast usually changes between acquisitions. Therefore it is necessary to implement a robust feature detection algorithm, which in turn produces landmarks that can be matched and used as the basis for non-rigid registration.

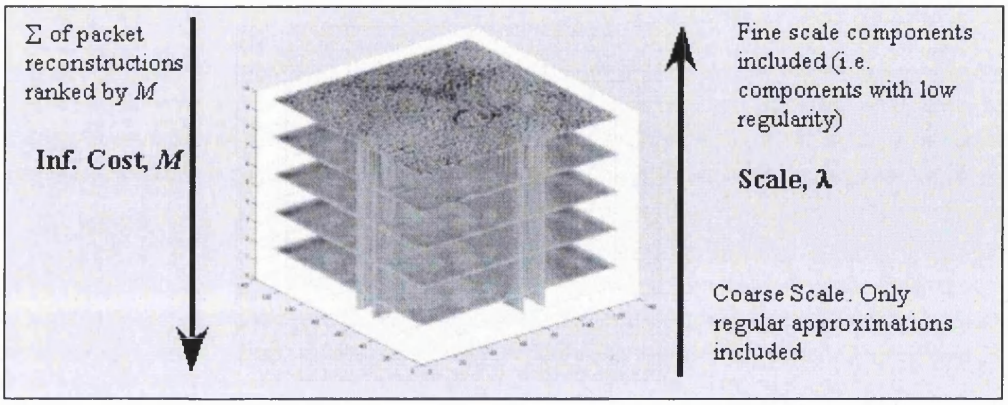
4.3.2 Segmentation of internal structures using wavelet-analysis

The algorithm used for segmenting internal structures in mammograms is described in [160] and is analytically presented in Appendix B. We have applied this multi-scale segmentation to mammogram pairs in order to detect regions of dense tissue that could be subsequently matched. The main steps of the algorithm are:

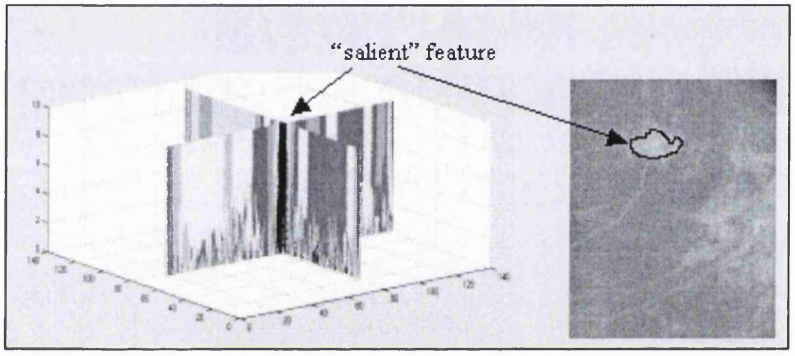
1. The mammogram pair is decomposed using the Coiflet wavelet packets. This particular wavelet was chosen because it yields good spatial localisation (e.g. it is edge preserving), has compact support and because it is effective even at detecting small regions such as microcalicifications.

2. After each mammogram is decomposed into a set of high-frequency and low-frequency images (with good spatial localisation of features), these are ranked by information content using an information cost function in the context of a "best basis" algorithm [110]. For this, an entropy measure is used and each wavelet subspace (filter superposition) is then cumulatively adjointly convolved, in order, with respect to the best-basis assessment of the decomposition. As is shown in Figure 25 (a), the result is a "stack" of reconstructions from minimum to maximum information content (dependent on the cost function). The top of the stack contains the fully reconstructed image while coarser features appear towards the bottom. This construction is used to track significant features through scale space and forms the basis of the feature segmentation.

3. After constructing the "stack", important features are detected as features that persist through the scale-space. This is done through region growing from the lowest scale towards the highest. A merging operator tracks the feature to the highest scale so that each feature can be represented with more "detailed" information. Figure 25 (b) is a representation of feature detection through the scale-space stack. In a mammogram pair, we track the n most important regions (usually n<=5) and subsequently we match them as described in the next section. More analytical segmentation results are presented in the example in section 4.5.



(a)



(b)

Figure 25: (a): The scale space "stack" for robust detection of mammogram structures, (b): A schematic representation of the detection of features through the scale-space (the stack is inverted to show the detection from coarse to fine).

4.4 Landmark matching and registration refinement

4.4.1 Matching internal structures

Based on the regions that are detected using the scale-space segmentation approach described in the previous sections, a set of internal landmarks is defined by a matching algorithm that includes the partial transformation (induced by the boundary alignment) in conjunction with scale, size and area information of the candidate matches. In the registration process the segmented regions are represented by their centroids. Due to the

small number of the segmented significant regions, matching becomes an easy task as we can introduce some further constraints that further reduce the search-space:

The initial search-space for a match in the first image is defined as a window in the second image whose size is proportional to the amount of displacement of the transformed internal landmarks using the boundary transformation. This is used to limit the possible matches to a "window" or neighbourhood. All the feature parameters (size, scale, relative motion) are used to drive a simple spatial searching algorithm that is conceptually similar to the iterated closest point method (ICP) [96-98]. Essentially these criteria are used as the basis of a "match rejection" filter. After evaluating a distance measure between landmarks to ensure that landmarks classified as a "match" have similar spatial properties and have demonstrated a change in geometrical correspondence as a result of the boundary deformation.

In the wavelet feature detection and subsequent feature matching, the chosen criteria (size, scale, relative motion) limit the error in search for feature correspondence. In particular, we restrict matching on the basis of features with similar scale localisation in the wavelet "feature stack". This effectively means that we are not interested in matching features that significantly change size (due to differences in breast compression, for example).

On average, depending on the degree to which the breast is involuted we define 3-5 internal landmarks at the centers of mass of the corresponding wavelet-defined regions. In most cases, as shown in Figures 26 (a) and (b), the segmented salient regions in the partially warped and the target image move towards each other and exhibit a significant degree of overlap facilitating the matching process.

As mentioned before, the area of a region is used as a matching parameter. The objective is to ensure that spatially localised features are not matched to larger features. In addition, the relative motion between the landmarks in the undeformed and partially registered (via boundary curvature) data adds significant constraints to the matching process. This is

illustrated in Figure 26 (c), where the relative motion of landmarks is included as a weighting factor in the match. Our matching algorithm is conceptually similar to Kok-Wiles's work [100], where topological constraints are used as a matching criterion in order to reduce the search space.

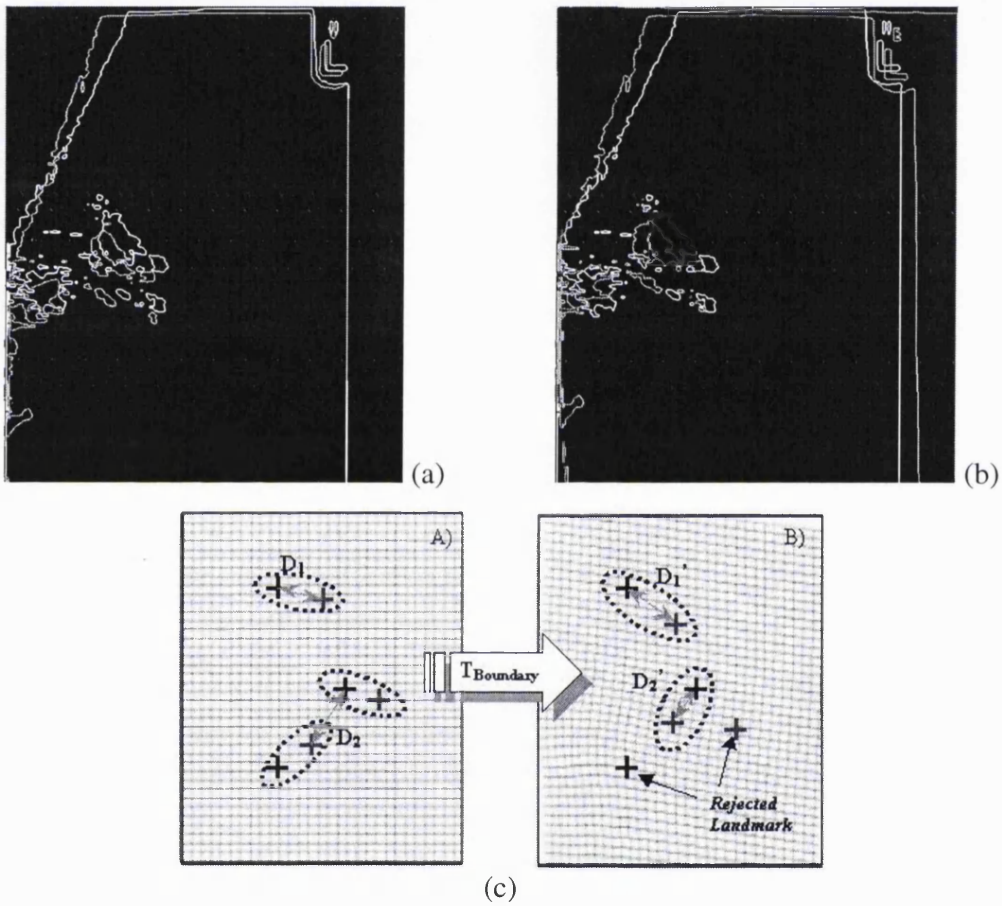


Figure 26: (a): Segmented "significant" regions of a temporal pair superimposed. (b) The same between the "target" and the partially registered image. The corresponding regions "move" towards each other as the partial transformation approximates the motion under different compression. (c): The boundary registration is included as rejection criterion for the matching algorithm. Before boundary registration (A) there are two candidate pairs with distances D_1 and D_2. After the boundary registration (B), D_1' is maintained, while the rest of the closest matches changed according to the deformation.

4.4.2 Final registration using an approximation scheme

4.4.2.1 Overview

The last step in the registration process is to include both the boundary (curvature-based) landmarks and the internal landmarks (as outlined in Figure 1). However, at this stage of the registration process, an approximation (rather than an interpolation) scheme is used to compute the elastic deformation. This is to account for possible inaccuracies in landmark representation, as well as to produce a smooth deformation that takes into consideration the relative importance of the matched regions (represented by their center of mass). It is worth mentioning that in a clinical deployment of this algorithm, the user should confirm that the detected internal landmarks have a good anatomical correspondence. Due to the large variability in tissue architecture, the anatomical significance of the detected internal structures is very hard to assess. However, the segmentation method discussed previously detects the dominant glandular tissue structures in each mammogram. Next, we discuss the approximation scheme used. The aimed is to account for representing the matched regions only by a centroid and the overall uncertainty in landmark localisation (e.g. localisation of very small regions is prone to a more significant error than to large regions).

4.4.2.2 Approximating Radial Basis Functions

The boundary points and the internal landmarks (computed by the wavelet analysis) together control a thin-plate spline approximation technique, which gives the final registration. Since the confidence in landmark detection is higher for the boundary landmarks, we need to employ an approximation scheme that allows us to control the smoothness and the individual weighting for each landmark. For example, a pair of internal landmarks that correspond to "large areas" of tissue should play a more important role than a smaller one.

To implement this in a registration algorithm, in such a way that landmark localisation errors can be accommodated, it is necessary to weaken the interpolation condition in such

a way that the resulting transformation is a compromise between smoothness and data adaptation. The new functional $J_\lambda(\mathbf{t})$ to be minimised is a function of λ, a regularisation parameters that controls the balance between the smoothness and the approximation of the transformation [79]:

$$J_\lambda(\mathbf{t}) = \sum_{i=1}^{n} \frac{|\mathbf{q}_i - \mathbf{p}_i|^2}{\sigma_i^2} + \lambda \cdot J_2^2(\mathbf{t})$$

Where $\mathbf{q}_i$ and $\mathbf{p}_i$ are the landmark pairs, $J_2(\mathbf{t})$ is the bending energy of the transformation $\mathbf{t}$ and λ is a regularisation parameter that controls the trade-off between the smoothness of the transformation and adaptation to the local transformations induced by the data. For $\lambda \to 0$ the solution is a very good approximation of the data (we trust the data to be accurate), while for $\lambda \to \infty$ we obtain a very smooth transformation with very little adaptation to local deformations. If we are able to somehow "predict" the accuracy of our landmarks, we can weight each landmark pair by the inverse of the variance σ_i^2. A set of weights (the inverse of the variance σ_i^2) needs to be defined in order to characterise the uncertainty in the localisation of the landmarks. After matching the segmented structures the "goodness" of each match is characterised by the persistance of the regions in the wavelet stack (i.e. the "volume" of the region *through* the scale-space stack [160]). These scores are used (rescaled between 1 and 100 and the higher value is assigned to the boundary landmarks) as the landmark weights in the approximation scheme. The global transformation smoothly transforms the boundary while local deformations occur on the inside according to the importance of the internal matches. The number of internal landmarks detected depends on the number of significant regions that propagate in the scale-space stack and are consistent in both images in the pair.

The scale localisation information is used to control the σ_i^2 terms that have the effect of weighting the smoothness in the deformation by the saliency of the landmark. In this way, greater deformation influence can be attributed to features that have stronger spatial localisation. The computational implementation is similar to the interpolation problem since only the K submatrix from the interpolation-implementation needs to be changed:

$$K' = K + \lambda \cdot V^{-1}$$

where $V = diag\{1/\sigma_1^2, \ldots, 1/\sigma_n^2\}$.

Figure 27 illustrates the difference in the resulting transformation, if an approximation scheme is used instead of interpolation. In Figure 27 (a) all the landmarks are fixed in the same position, except the middle one that is significantly displayed. Figure 27 (b) is the interpolation result, where every single landmark is forced to move exactly to the new position. By changing the overall smoothness (via approximation) we obtain the result of image 27 (c) even if the landmark co-ordinates remain exactly the same as before. By weighting less the largely displaced landmark (that could be the case of a "low confidence" match) we reduce its effect on the final transformation as is illustrated in Figure 27 (d).

In the next section, we show analytically an example of the proposed registration method. We also discuss ways to validate the non-rigid registration and assess the improvement in registration when internal landmarks are used.

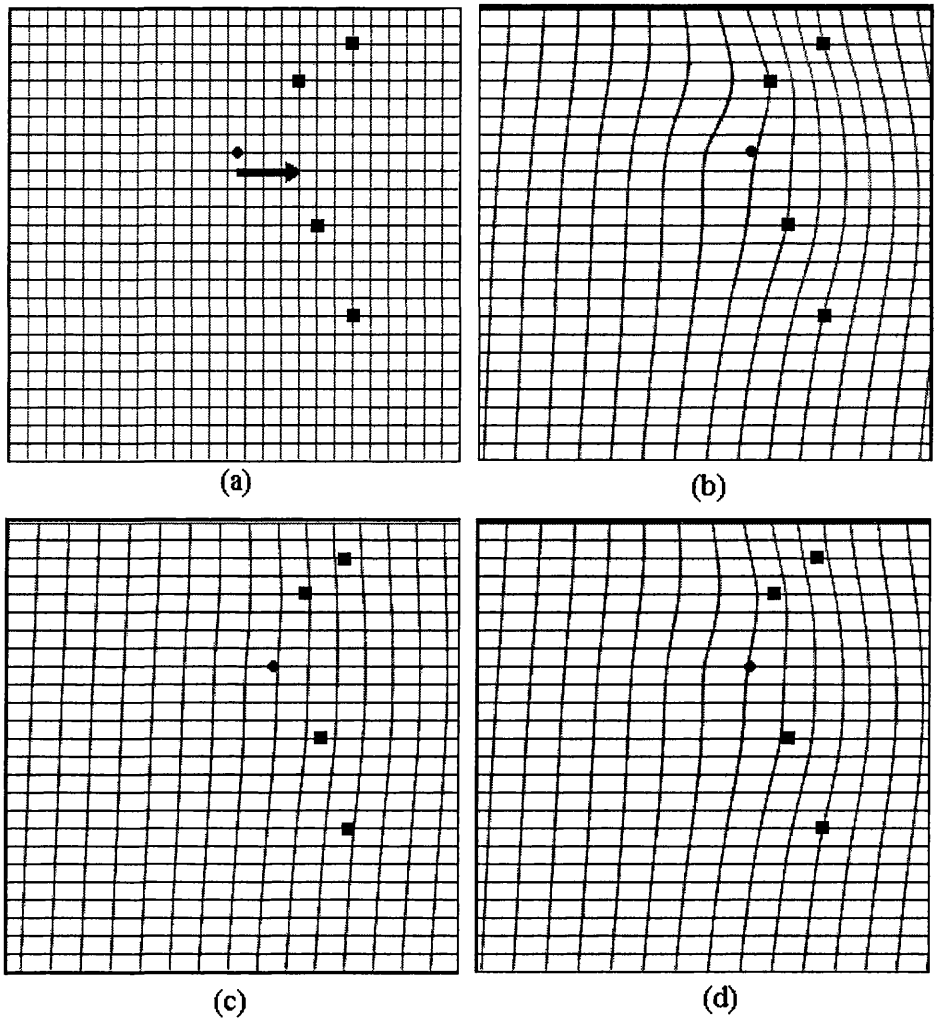


Figure 27: (a) Assuming that all the square landmarks remain in the same position while the central one (drawn in circle) has to move significantly as a result of a matching calculation. In (b): Registration is performed using thin-plate spline interpolation and the transformation moves the landmarks to the exact corresponding locations. In (c): the approximation scheme discussed above is employed and a "smoother" transformation that approximates the landmark requirements is obtained. In (d): the largely displaced landmark (the circle one) is weighted less than all the others and that reduces its effect (comparing to (b)) in the resulting deformation field.

4.5 Results and validation

4.5.1 Overview

Using the techniques described in this chapter 50 pairs of mammograms have been registered. A significant number of this data set is included in a separate ***gallery*** in Appendix E of the thesis (we present a typical sample since 2 pages are needed to present each pair). For these examples we show the registration results and, in addition, clinical information, comments on the registration results, the difference images, the transformation grids and in some cases the joint histograms have been included. A summary of the performance on all 50 mammogram pairs is given in Table 3.

In this section we describe how the proposed method was tested:

- The limitations in evaluating non-rigid registration in mammograms are discussed in section 4.5.2.

- Qualitative results on the registered data set are included in section 4.5.3

- Assessment of the improvement in registration using internal landmarks (with an example of registration including the calculation of internal landmarks) is presented in section 4.5.4

- In section 4.5.5, the accuracy in registration is assessed by comparing the calculated registration results with "obvious" landmarks marked by the user.

4.5.2 Limitations in evaluation

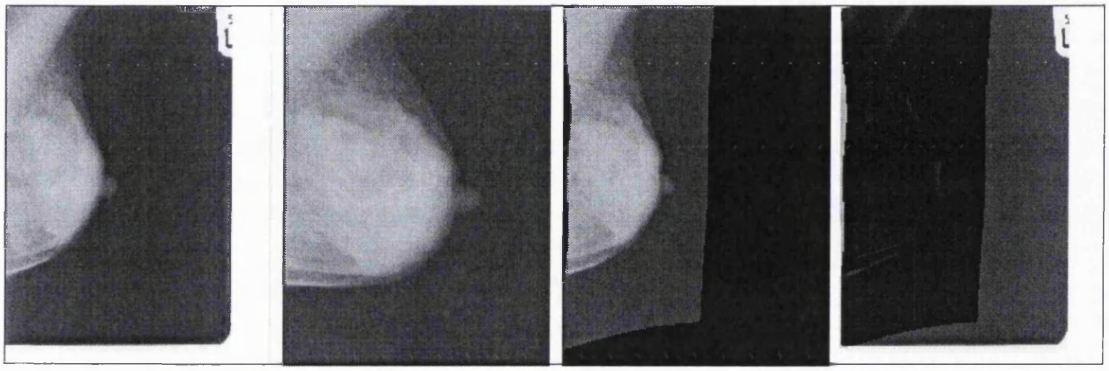
As was discussed in section 3.9, it is very hard to evaluate non-rigid registration techniques since it is difficult to establish ground truth. This is particularly truth in the case of mammogram registration since there are actual temporal changes in the

mammographic appearance of the breast. As a consequence, even if the registration method is perfect, the aligned images are not expected to be *identical.* This means that we can only rely on a combination of subjective criteria and image characteristics to assess the improvement in similarity between the mammogram pair. In addition, it is impossible to use fiducials inside the breast which could provide a ground truth for registration.

Most researchers use "artificial data" to evaluate such complex registration problems. An example is illustrated in Figure 28 (a), where we apply a random transform (taken from a mammogram pair so that it reflects a *realistic* breast deformation) to Figure 28 (b) thus obtaining the deformed mammogram shown in Figure 28 (b). By applying our registration image the original mammogram is recovered as shown in Figure 28 (c) (warped image) and Figure 28 (d) (difference image after registration). This kind of experiments offer a ground truth for evaluation (the images have to be identical). However, in such an experiment we can't mimic the actual anatomical changes that occur in the breast during acquisitions and are amplified by differences in compression and imaging parameters. For this reason we use *real temporal data* in order to validate the proposed registration technique, even if this makes establishing a ground truth a more difficult task.

First, we qualitatively characterise the registration results in section 4.5.3 while in section 4.5.4 an example of registration is presented, in order to demonstrate the improvement of the result when internal landmarks are used. This improvement is also assessed by the clinician in 25 mammogram pairs from the data set. To establish a more quantitative evaluation, in section 4.5.5 we propose the use of marked, "obvious" landmarks in the mammogram sequences.

Finally, as mentioned in the introduction a large number of restoration examples is included in the Appendix E of the thesis, so that the reader can better understand the application of the suggested method in a large number of mammogram pairs.



(a) (b) (c) (d)

Figure 28: Creating an artificial experiment for evaluating registration, (a) The original mammogram, (b): A random mammogram transform is applied and the mammogram is significantly deformed non-rigidly, (c): Using the mammogram registration technique the mammogram is aligned to its original shape, (d): The difference image shows that the images are almost identical, however this artificial experiment does not represent the actual mammgraphic changes encountered in real temporal data.

4.5.3 Qualitative characterisation of the registered images

In this section we provide a short description of the registration results in the 50 images that were registered using the registration method presented in this chapter. As has been discussed before, the registration process is not expected to make the mammogram pair look identical since there are actual temporal changes in the breast structure and thus in the mammographic appearance. The success of the registration can be evaluated in terms of the increased similarity in image features after geometrical alignment. The viewer characterised the geometrical alignment based on observing the alignment of the breast boundary and of the internal breast tissues. The criteria for assessing the quality of registeration were:

- Good: If the images are well registered and no misregistration "shadows"appear in the difference image after registration.

- Average: If the registration result looks satisfactory but at the same time there are regions of poor overlap in the difference image resulting in "shadows".

- Poor: If the features (boundary or internal strucures) are not aligned well (e.g. the nipple in the boundary is in a different location in the registered images) resulting in a large number of misregistration "shadows" in the difference image.

In these examples, the full registration process is characterised, while in the next section we assess the impovement of the registration result after the internal landmarks are included.

These observations are summarise in Table 3:

Boundary alignment	Internal correspondence
Good: 100%	Good: 70%
Average: 0%	Average: 25%
Poor: 0%	Poor: 5%

Table 3: Comments on registration results in 50 mammogram pairs: The viewer classified the results in three categories according to the alignment of image features.

4.5.4 Assessing the improvement in registration

In this section, a complete registration example is illustrated, including images of the intermediate steps. The aim of this example is to demonstrate how the inclusion of internal landmarks can better approximate the deformation of the breast due to different

degrees of compression, leading to an improved alignment of the images. A small number of internal correspondences can significantly improve the registration result since it can account for tissue motion inside the breast under compression. To assess the improvement in registration we suggest the following:

- Calculation of the difference image after registration. This provides a visual verification that the images have been aligned (the misregistration "shadows" are significantly reduced after registration).

- The joint histogram is another indication that the image information context has been aligned compared to before registration. Before registration the values of the histogram are spread out while after the registration they are more concentrated around the main diagonal.

- Assessment of the improvement by the clinician; the clinician decides if inclusion of internal landmarks significantly improves the quality of the correspondence between the mammogram pair.

In Figure 29 (a) and (b), a temporal pair of mammogram is registered, first using just boundary points (Figure 29(c) is the transformed 29 (a)). Images 29 (d) and (e) show the matched regions based on the multi-scale wavelet analysis, superimposed onto images 29 (c) and (b). Using the centers of mass of the corresponing regions, we re-register the images as described in section 4.4.2. Figure 29 (f), is the resulting warped image.

The last three images show the difference image after the applying the proposed registration (in Figure 29 (i)); it clearly shows the improvement in the registration (compared both to no registration and to the boundary registration difference images in Figures 29 (g) and (h)). Note that all the important features overlap, and the "shadows" corresponding to misregistations in Figures 29 (g) and (h) disappear. The deformation grid, shown in Figure 30, is an alternative way to visualise the partial registration (Figure 30 (b)) and the improved registration (Figure 30 (c)).

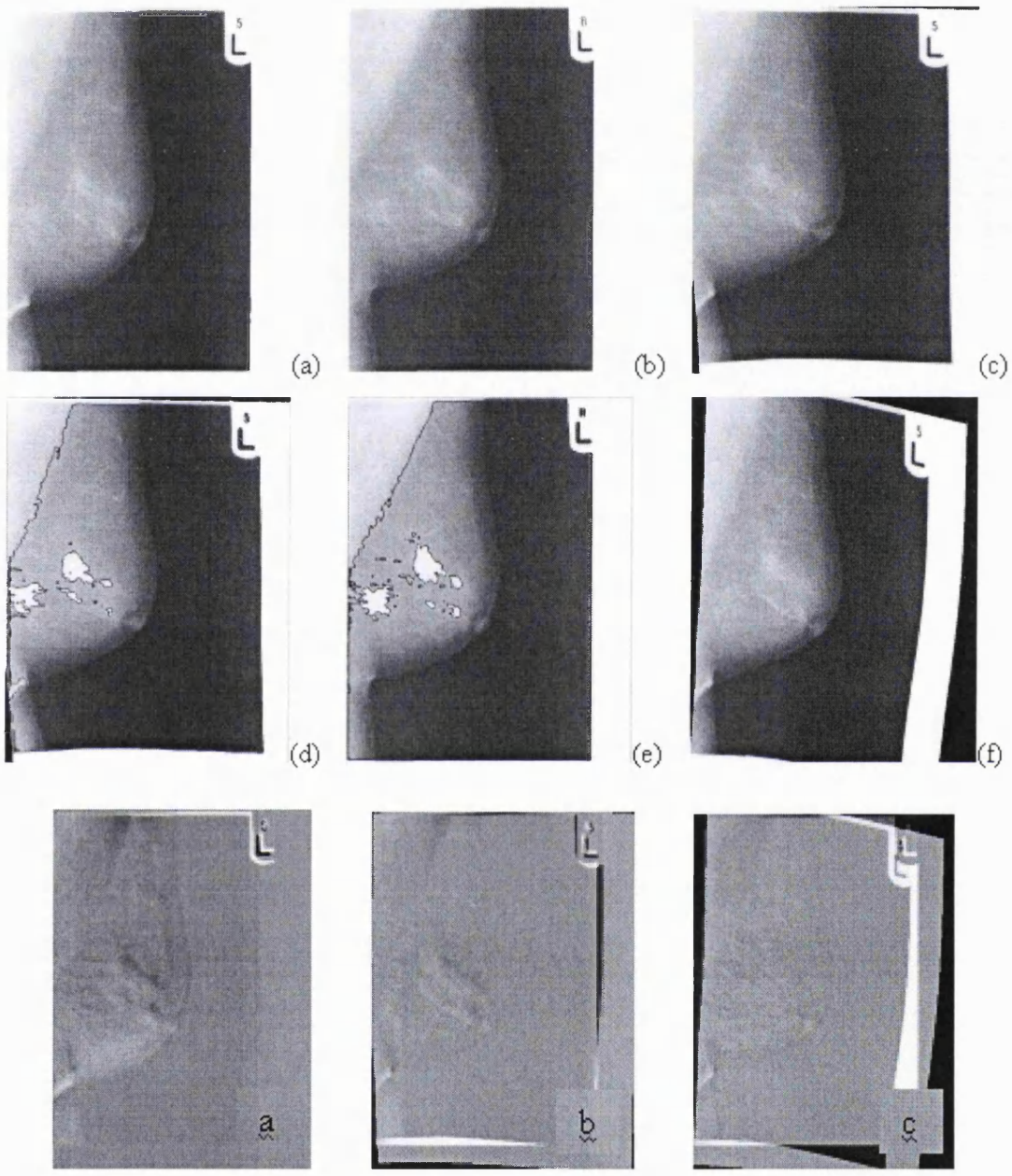


Figure 28: (a) and (b): A temporal set of images, (c) Image (a) warped in the co-ordinates of image (b) using the boundary landmarks. (d) and (e): The matched regions based on the multi-scale wavelet analysis, superimposed to images (c) and (b), (f): The warped (a) to (b) after the final registration. Difference images: (g): Before registration, (h): After partial registration and (i): After final registration including internal landmarks.

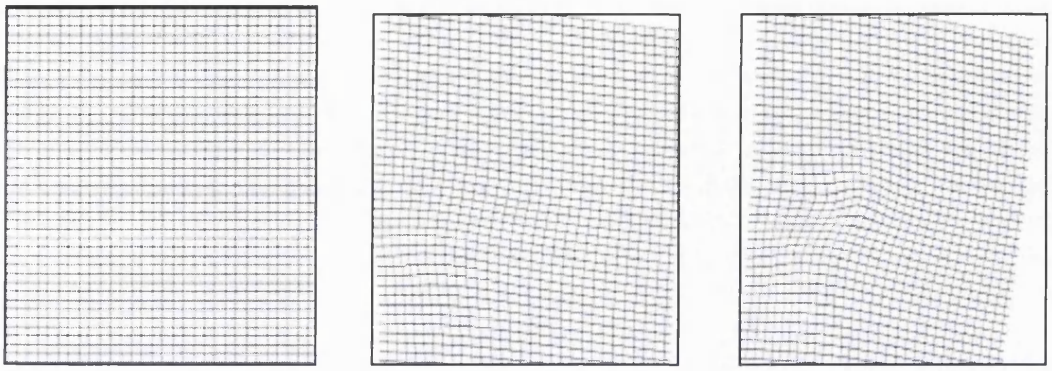


Figure 30: (a): The orthogonal grid, (b): The deformed grid for the partial transformation is relatively smooth inside the breast, (c): After the final registration the deformed grid depicts the relatively non-smooth motion inside the breast under changing compression more realistically.

In a clinical assessment of the technique, an experienced radiologist judged that the correspondence between the registered images was improved in 20 out of 25 pairs of temporal mammograms after the inclusion of internal landmarks. This is in agreement with the difference image (less or no misregistration "shadows") and joint distribuition histogram (intensity pair values concentrated more towards the main diagonal).

In Table 4, the clinical assessement concerning the improvent in registration with tne inclusion of internal landmarks is presented together with the ruduction in the standard deviation of the difference image after registration in the same mammogram pairs. This can be seen as a numerical interpretation of Figure 29; using only the boundary makes most of the shadows in the difference image dissapear and this effect is amplified when internal structures are included in the registration process as well as the boundary. The cases of "no improvement" in Table 4, represent the mammogram pairs where the tissue was predominantly fat (involuted) and as a result no bright regions could be segmented and matched.

Improvement of the registration result when including internal landmarks	Reduction in the standard deviation of the difference image
Significant: in 20 mammogram pairs	10%-35%
Not significant: in 5 mammogram pairs	0%-15%

Table 4: Clinical assessment of the improvement in registration using internal landmarks in 25 mammogram pairs and comparison with the reduction in the standard deviation of the difference image after geometrical alignment for the same cases.

The success of this method is limited by the degree of involution in the breast tissue. The more involuted the breast, the less "significant" internal structures we can detect using our method. However, the 2-stage nature of the process ensures that at least the breast boundary can be aligned.

4.5.5 Assessing the accuracy of registration

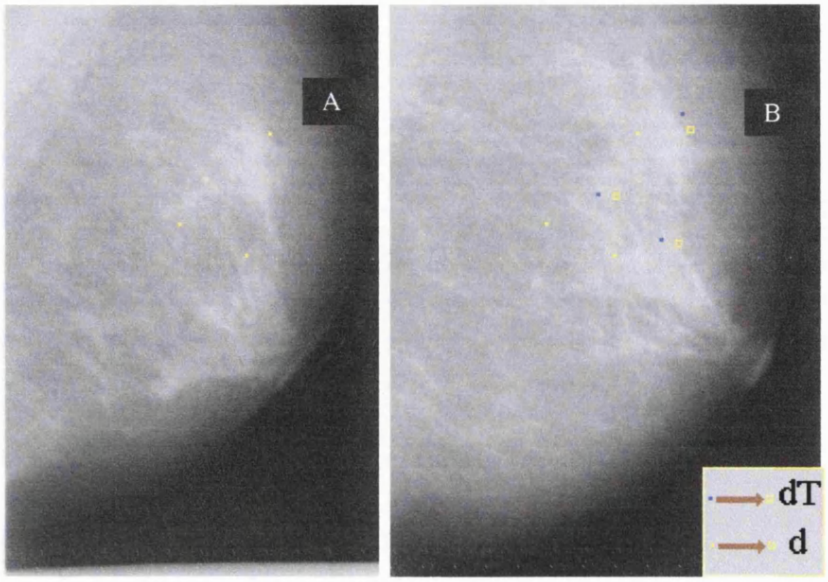
Assessing the accuracy of a non-rigid registration algorithm is a very difficult task (as was commented in chapter 3). There are no globally accepted "standards". In addition, unlike cases like brain imaging, significant fiducials can't be used in the breast, (except maybe the possibility to use skin markers but even then that could only be useful for a more accurate detection of the nipple).

In this section we suggest a simple method for assessing the accuracy of our technique. Temporal mammograms often exhibit several "obvious" correspondences since the tissue architecture is usually preserved. This experiment uses 10 pairs of mammograms where the architecture was preserved so that "obvious" correspondences could be used as

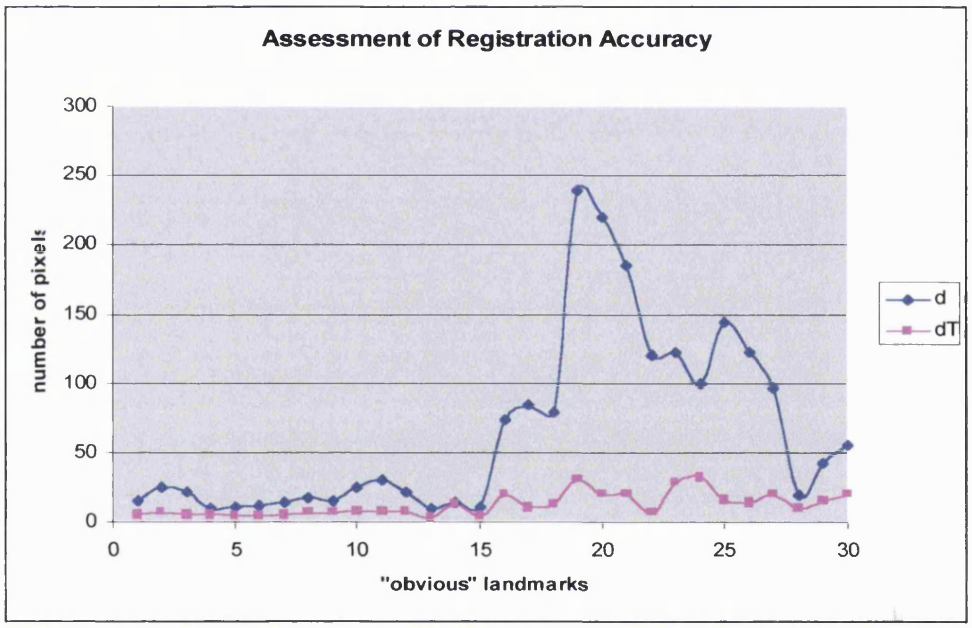
"ground truth". In each pair, three "obvious" landmarks were selected, as shown in Figure 31. These landmarks were most often regions of dense tissue or benign masses. In addition, "obvious" landmarks had to be significantly different (>100 pixels in 300 micron images) from the landmarks used for the registration in order to assess the *global* improvement in correspondence. Finally, half of the mammogram pairs had a very good initial correspondence (points 0 to 15 in Figure 31 (b)) in order to observe the improvement in cases where the displacement due to registration is small.

An example is shown in Figure 31. The landmarks are shown in yellow in the source image (A in Figure 31 (a)), and as yellow squares in the target image (B in Figure 31 (a)). The absolute distance between the corresponding points, **d**, before any registration describes the original error before registration. By applying the calculated transformation to the source image, the three points are mapped to the corresponding ones (shown in blue), in the target image. The distance between the "transformed" points (shown in blue) and the reference points (yellow squares) in the target image, **dT**, represents the error after registration. In addition, the comparison of **d** with **dT** is a measure of improvement of the correspondence of features after registration.

Figure 31 (b), shows the error in correspondence of the "obvious" landmarks before and after registration. It is obvious that registering the images improves the correspondance of tissue structures in mammogram pairs. In a clinical context, the estimated transformation (for alignment of mammogram pairs), could be used to estimate the location of points or regions in temporal mammogram sequences, making the comparison of mammograms an easier task.



(a)



(b)

Figure 31: (a) The user selected 3 anatomically significant points in (A) and by using our registration method we estimate their locations in (B), shown with blue dots. If the clinician (or the user) marks the corresponding points in the second image (yellow squares) the error distance dT can be used to validate the accuracy of the transformation. (b): The registration error in pixels for pre-selected "obvious" landmarks. Before registration d represents the error without registration, while dT is the error after alignment.

4.6 Registration issues and applications

4.6.1 Bilateral mammograms

As mentioned both in this chapter and in the previous one, most efforts on mammogram registration to date have concentrated on bilateral mammograms. Although this is not the main concern of this thesis, the method described can be used to register bilateral mammograms and help the clinician determine regions of significant asymmetry. The boundary can always be used to register bilateral mammograms, independently of the degree of "architectural similarity" between them.

The wavelet segmentation method has the potential also to be used to match bilateral mammograms pairs providing that a "structural" similarity exists. However, we have found that in many cases the architectural similarity between the breasts is not strong, and internal correspondences are difficult to establish. In Figure 32, there is an obvious correspondence between the segmented (based on the wavelet analysis) regions of the bilateral pair.

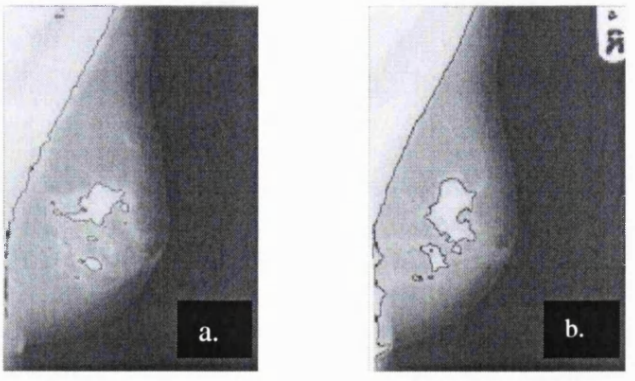


Figure 32: (a), (b): A bilateral pair of mammograms with the segmented "significant" regions superimposed. This clear correspondance of significant features is not always the case in bilateral mammograms.

Bilateral registration can also be used in the case of women that follow Hormone Replacement Therapy (HRT) as in some cases the differential response to oestrogens can be a sign of an underlying malignancy (a more "active" breast is more likely to develop a tumour). However, we emphasise that the architectural similarity in bilateral mammograms is poorer than the temporal mammogram case, and it's very difficult to ensure consistency in the clinical importance of the results.

In Figure 33 an interesting example of bilateral registration is illustrated, where the registration result maximises the perceived similarity of the two images thus facilitating comparison.

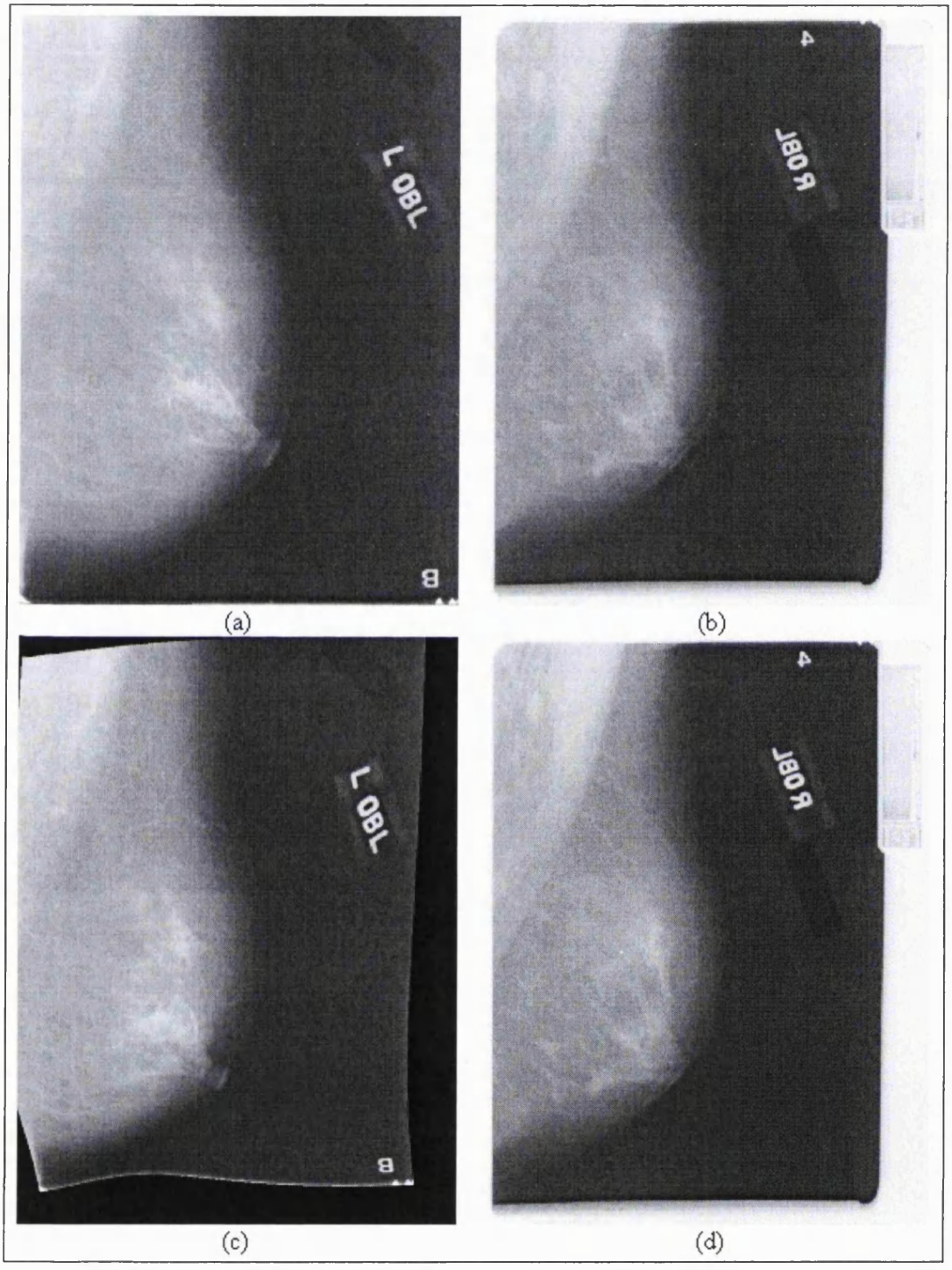


Figure 33: Registration of bilateral MLO views. (a):The original left MLO mammogram, (b): The original right MLO mammogram, (c): The left MLO registered to the Right MLO and (d): again the left MLO for easier comparison with (c)

4.6.2 The role of the pectoral muscle

The pectoral muscle moves independently of the rest of the breast during the X-ray acquisition. The area of the pectoral muscle is also variable between bilateral and temporal mammograms of the same patient. This can give rise to several issues of accuracy and consistency concerning registration or quantification techniques.

Figure 34, illustrates that registration using the pectoral muscle can be severely in error, especially if landmarks from the inside of the breast have not been included in the interpolation scheme. Since the pectoral muscle is currently being used in mammogram registration ([107], [108]), this is an important issue to address.

As illustrated in Figure 34, trying to align the pectoral muscle without including landmarks from the inside of the breast (as in [107]), leads to an increase in the distance between corresponding features. In fact, the difference image 34 (f) reveals that although the pectoral muscles are aligned, the fibroadenoma (visible in both images) is *further* away than in the unregistered images. We have observed from differential compression mammography [4], that the pectoral muscle seems to move independently of the breast tissue under different compression and this might be a possible explanation. For that reason, further investigation is necessary to clarify whether or not the pectoral muscle should be taken into consideration in mammogram registration

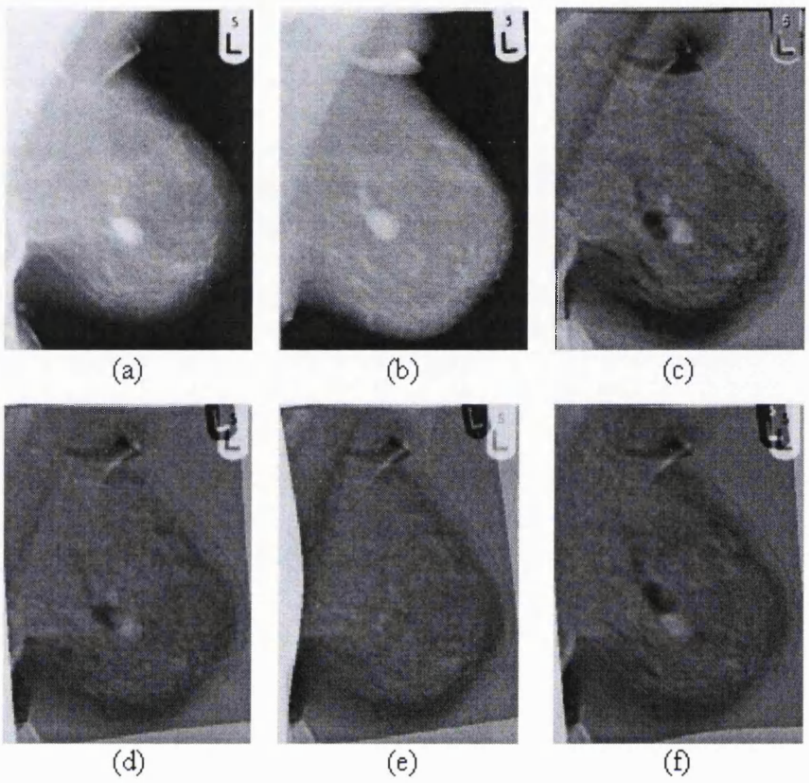


Figure 34: (a),(b): A temporal pair of mammograms with a fibroadenoma present, (c): Difference image before registration, (d): After registration using five points in the boundary the difference image indicates that we have compensated for some (but not all) of the tissue deformations, (e): By manually adding two internal landmarks, registration is improved, (f): If we use two points from the line of the pectoral muscle instead, although we align the pectoral muscle, the mass is further displaced than in the unregistered difference image (c).

4.6.3 Registration and interval cancers

We consider mammogram registration to be an important element of future clinical mammography software, especially since digital mammography will allow such algorithms to be incorporated in a faster and more efficient manner. Registering the most recent mammogram taken with the previous ones can provide information about regions of significant difference and assist the clinician in detecting subtle new changes.

Figure 35 (a) and (b) shows an example of an interval cancer. Figure 35 (c) is the second image warped in the co-ordinates of the first, while 35 (d) shows the difference image after registration. The difference image can be used for the automatic detection and segmentation of the cancer using a one-tailed t-test on the pixel intensities inside the pre-calculated breast boundary (Figure 35 (e)). In that way we avoid segmenting the line of "misregistration" along the boundary. Since the segmented cancer has been transformed to the co-ordinates of the first mammogram we can estimate the region of future development of cancer in the first, "normal" mammogram by superimposing the boundary of the transformed cancer region (Figure 35 (f)).

As mentioned in the beginning of this thesis, a significant number of "interval cancers" are missed (and could have been diagnosed from a previous mammogram). Registration could be used for retrospective studies in order to examine the region of the cancer prior to diagnosis. This could help in determining the number of "true" interval cancers and therefore the incidence of breast cancer.

Although we believe that registration can reduce the number of "missed" cancers in a screening context, it is a very difficult hypothesis to verify. The first reason is that we need to test many more cases where the cancer was "missed" and retrospectively verified by the clinician that "it could be detected", and such data is difficult to compile, particularly as inter- and intra- observer variation is high. In addition, mammogram registration would have to be used in a test trial related to screening before all the expected benefits of such a technique are challenged in practice.

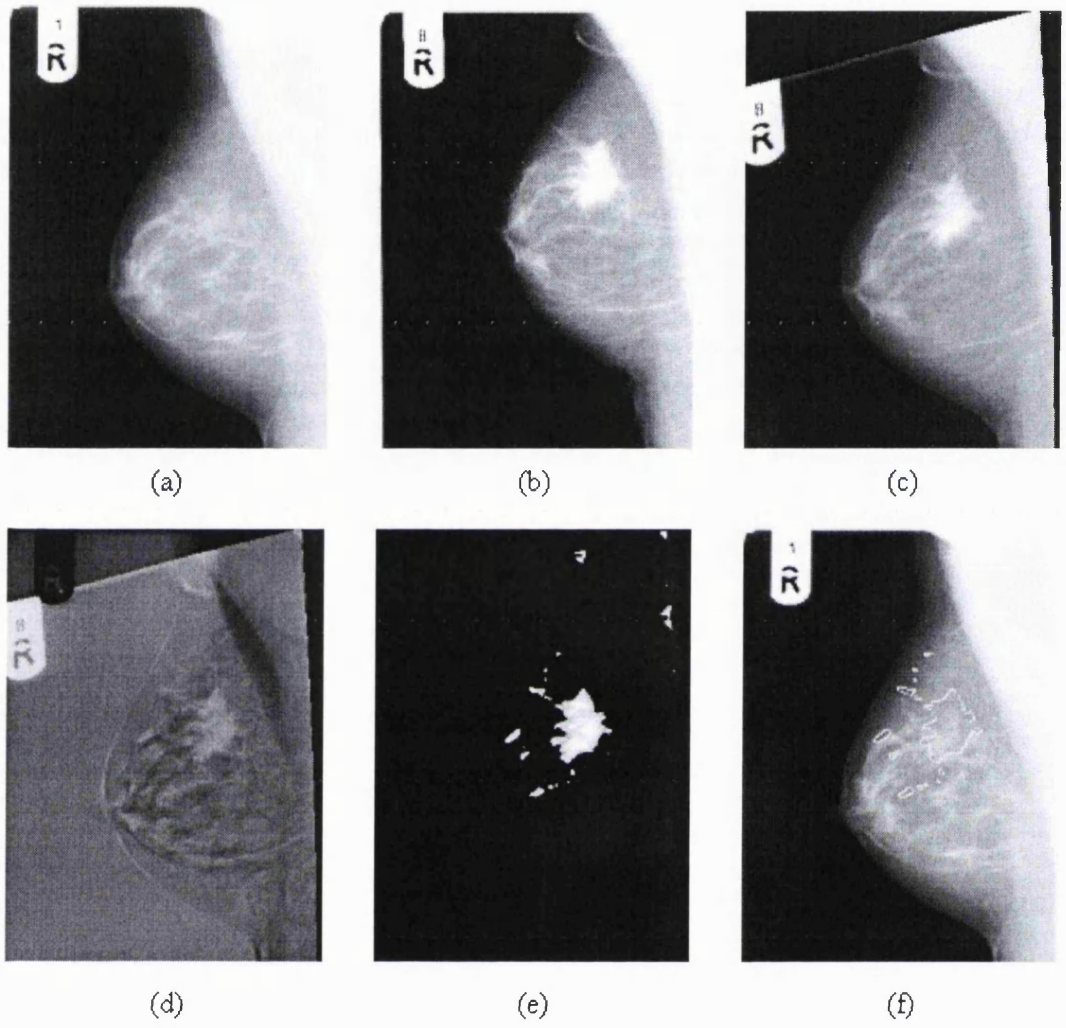


Figure 35: (a): Before the development and diagnosis of the cancer, (b): A clear cancer has developed, (c): Mammogram (b) warped in the co-ordinates of mammogram (a), (d) Difference image after registration, (e): Cancer segmented from (d), (f): The transformed boundary of the cancer superimposed on the "normal" mammogram (a), indicates the region of the future development of cancer.

4.7 Discussion and summary

In this chapter, we demonstrated a method that first aligns temporal mammograms based on the boundary. We demonstrated how the breast-boundary can be consistently used to

resister mammogram pairs and presented evaluation results concerning the anatomical significance of the three proposed maximum curvature landmarks in the breast-boundary (nipple, axilla and rib points).

Multi-scale analysis provides a reliable framework for establishing correspondences between significant regions inside the breast. In this way, the algorithm presented here is an improvement on that presented by [100, 102]. Using internal landmarks, the registration result is improved, as was asserted by the clinician and the difference images after registration as well as the joint histograms of the aligned images. Even though matching points inside the breast is difficult due to temporal changes and depends upon the extent to which the architecture (or topology of the surface) is preserved, the multi-scale segmentation method used, reliably locates regions of dense tissue that appear in both temporal mammograms. Additionally, using the thin-plate approximation scheme (section 4.4.2), we can weight our internal landmarks according to their size and scale and therefore compensate for landmark localisation errors.

Obvious landmarks, were used as ground truth for estimating the correspondence error before and after registration. After registration, the correspondence was improved from 65.2 average and 65.7 standard deviation (units in pixels) to 12.2 and 8.2 respectively. Since the images were digitised at 300 microns, this translates into an error reduction from 2cm before registration on average, to 0.37cm after registration. More experiments are necessary to establish the improvement in correspondence of features (since it depends on the initial correspondence of the images) and determine the accuracy that is necessary for clinical use.

Mammogram registration could be a useful tool for aiding the clinician to detect abnormalities by comparing a mammogram with a previous one, as significantly "different" regions can be detected in the difference image and are their location can be easily identified in both mammograms (after registration). In addition, we can use the same registration technique to align HRT sequences in order to detect overall changes in tissue composition and to be able to estimate the site of local changes in tissue density;

this work is discussed on the next chapter. Finally, Behrenbruch et al. [32, 33], demonstrated that this registration framework can be extended to achieve 3D-MRI / 2D-X-ray data fusion as is summarised in the concluding chapter of the thesis.

More generically, mammogram registration is important for future digital mammography systems, where the clinician will be able to examine both the original and the aligned temporal mammogram pairs when reading mammograms. In addition, other diagnostic tools (e.g. microcalcification detection, or image enhancement) could be used in conjunction with registration in order to compare features (or examine "prompted" regions) in the aligned mammogram pair. Temporal comparison is already an important part of mammogram reading in European screening programmes and computer aided mammogram interpretation tools (e.g. registration) are hoped to increase the early detection rate of breast cancer.

In temporal HRT mammograms, tissue quantification becomes a crucial issue (unlike non-HRT mammograms where involution is a physiological process that is not correlated with cancer). For this reason, we need to combine mammogram normalisation (using the h_{int} representation) and registration of temporal mammograms aiming to develop a robust temporal tissue-quantification framework. The issues of tissue quantification and registration of HRT sequences, is the subject of the next chapter.

4.8 Concluding statement

A mammogram registration method has been proposed in this chapter. The basic aim to develop a robust registration method was achieved. The set of landmarks used in that method comprise of a number of boundary landmarks which very often have an anatomical significance and a small number of internal landmarks that correspond to the regions that appear to be similar in the temporal pair. Thus the hypothesis that patient specific landmarks can be calculated was achieved. This set of points is enough to

calculate the geometrical transformation necessary to align the mammogram pair which confirms our working hypothesis that such a set of landmarks could be defined in each mammogram.

However, as was explained in section 4.5.2, the assessment of the quality of the non-rigid registration is a rather subjective task since the breast changes between acquisitions. Nevertheless, some validation experiments are performed in the same section in order to assess the quality (in terms of feature alignment) and the accuracy of the suggested method. In addition, the role of the pectoral muscle is investigated as well as the application of the suggested method for bilateral matching. Finally, a number of clinical applications is discussed (e.g. registration and interval cancers). In this thesis, we chose to analyse data from HRT users, in order to develop an image analysis framework for assessing the density changes, which are crucial. This work is reported in the next chapter

Chapter 5: Temporal analysis and breast tissue quantification of HRT sequences

5.1 Image Analysis for HRT mammograms

Introduction

In this chapter, we are concerned with women that use Hormone Replacement Therapy. In general, women choose to use HRT in order to relieve menopause-related symptoms or most importantly, prophylactically against osteoporosis and cardiovascular disease. Although research findings remain controversial, there is a concern that HRT can increase the risk of breast cancer especially for long-term users. A theoretical explanation could be that the hormone-stimulated breast cell proliferation increases the probability of cancer in HRT users. It is believed that invasive breast cancer evolves from normal epithelium of the terminal duct or lobular unit* through various abnormal proliferative lesions (e.g. typical and atypical hyperplasia). Female hormones (mainly oestrogen and progesterone) are involved in regulating the differentiation and proliferation of normal epithelial cells [163].

From an image analysis standpoint, HRT mammograms can be considered as a special case of temporal mammography, where the clinician is interested in changes over time. However, in non-users, the breast tissue gradually involutes into fat, a physiological process that favours diagnosis since a new growth will be easier to detect mammoraphically. In HRT-users, if the woman responds to the exogenous hormones, the reverse process can take place: glandular tissue may regenerate and this process should be carefully monitored by the clinician in order to diagnose an abnormality as early as possible. For these reasons we believe that, in conjunction with registration, there is a need to be able to *quantitatively compare* successive HRT mammograms. In other words, quantitative comparison can provide information about density changes, while registration can *help* the clinician estimate the location of these changes.

* Breast anatomy is summarised in Appendix A

In the next sections, we present the aims of our work, and we suggest how temporal registration and quantification could assist the clinician in assessing the risk of breast cancer for the *individual* woman.

The domain and the aims

The work in this chapter aims to address the Medical Image Analysis problems related to women that use Hormone Replacement Therapy (HRT) and to suggest a framework for analysing HRT sequences in terms of mammogram density. More specifically the aims of our work on HRT mammograms are:

- As was mentioned in the previous section (clinical motivation) our first aim is to provide a detailed literature review in order to understand the issues related to the use of HRT. Due to the controversy concerning the increased risk of breast cancer, it is important to present a detailed review of the medical problem and various reported clinical results concerning the effects of HRT on the risk of cancer, breast density and mammographic sensitivity.

- Based on previous work on mammogram normalisation we aim to propose and validate *quantitative measures* for breast density. It is believed this is of great importance for mammography since identifying regions of local increase in density due to the use of HRT could potentially assist a clinician in the early detection of cancers. However, the lack of adequate and objective quantitative measures impedes serious examination of this conjecture. Here, our first aim is to propose and validate a set of quantitative measures on a global basis (i.e. calculate density measures using the entire mammogram image) by comparing them to the expert's description of density change and to previously proposed descriptors of density. Then we propose an **image processing framework** that could potentially allow the clinician to assess the breast density changes *on a local basis* and therefore better understand the risk of cancer due to HRT for the *individual woman.*

- More specifically, in this chapter we aim to present the basic components of a system that will allow the automatic quantitative comparison of temporal mammograms. This would require a combination of registration and quantification, so that the clinician could be informed about the temporal changes in density *on a local basis*. However, in this chapter we only aim to analyse the possible problems that could arise from combining registration and normalisation and not to test such a system in clinical practice.

The working hypothesis

The main hypothesis in this chapter is that a normalised mammogram representation (e.g. the h_{int} representation) can provide the basis for a quantitative comparison of mammogram density. This hypothesis is made because the h_{int} representation calculates the height of fatty tissue in each pixel thus providing an anatomical representation of the breast tissue. This is especially important when temporal mammogram pairs are considered since the variation in imaging parameters complicates the calculation of invariant quantitative measures of density.

More specifically, in this chapter we test the hypothesis that quantitative measures of tissue density based on the h_{int} representation can automatically characterise the actual changes in breast density over time. Equivalently, our hypothesis is that h_{int} changes of "interesting tissue" (i.e. non-fatty tissue) in the breast, over time can describe the actual changes in the fibroglandular tissue content of the breast.

Finally, in this chapter we suggest a method for comparing mammogram density on a local basis. Although we discuss the technical requirements and limitations of such a quantification method, its effectiveness in early detection of breast cancer is not tested in this thesis since it will necessitate significant number of clinical trials. However, this is a hypothesis based on clinical experience and is discussed in section 5.5, while in section 5.6 we analyse the possible clinical impact of automated density quantification (on a

global and local basis) on the clinical management of HRT users. The quantification methods used in this chapter are now discussed.

Description of the selected methodology and the experiments performed

In the first sections of this chapter, we discuss in detail the medical problems related to the use of HRT and present a literature review of the effects of HRT in tissue density and patterns, mammographic sensitivity and breast cancer (sections 5.1 to 5.4). This is in order to understand the controversy in the literature over the benefits of HRT and the requirements in terms of Medical Image Analysis that would provide clinicians with useful tools for assessing the risk of cancer for the individual woman who uses HRT.

The measures we develop are based on the h_{int} representation of interesting tissue. The reason for this is that it is the most reliable method proposed so far for intensity normalisation of mammogram pairs. For comparison, we also calculate the percentage of the area of glandular tissue in the mammogram, since it has been used in the past to characterise tissue density. These measures are discussed in section 5.7. We used 59 pairs of mammograms to validate the performance of each measure individually on describing temporal changes in tissue density. These changes are:

Tissue regeneration: If the breast density in the most recent mammogram is increased as a result of HRT.

Involution: If the breast density is decreased in the most recent mammogram. This is a physiological process for postmenopausal women and in the case of HRT users is a result of poor hormonal receptivity.

No change: In many cases no obvious change in tissue density is observed. Although it is not such an important case as tissue regeneration, it should be included in order to validate the suggested quantitative measures.

The ground truth is established by the clinician who examined the temporal mammogram pairs and classified the observed changes in the 3 categories mentioned above. The design of the global quantification experiment and the results are presented in section 5.8.

For assessing local changes a combined registration and image normalisation framework is proposed in section 5.9. However, when registering a pair of mammograms, it is often the case that the image extent is reduced (or increased) due to the necessary rescaling for the image alignment. For this reason a method that preserves the image "volume"(defined as the sum of the intensity or h_{int} values over the image) after registration is presented in section 5.10. In the same section a validation experiment of this method is presented.

Nevertheless, there are certain limitations in the design and evaluation of experiments relating global or local quantitative measures. These limitations, as well as a discussion about alternative methods that could test the suggested density measures (local and global) are discussed in section 5.11.

5.2 Medical Background

5.2.1 Typical hormonal changes at menopause

Today's life expectancy of more than 80 years means that menopause is a much more important life event, since the average woman can be postmenopausal for up to a third of her life. "Menopause" is a word that encapsulates the totality of the hormonal changes that occur in the years leading up to, and beyond, the final menstrual period of a woman. For some women, the hormonal deficiency related to menopause can lead to significant pathology. In others, the menopause induces severe discomfort and pain, which lead women to seek medical help.

Before the menopause, during each menstrual cycle, the levels of oestrogen increase at the beginning of the cycle, stimulating the growth of an egg in the ovary. Ovulation occurs at mid-cycle, and soon after oestrogen, together with progesterone, stimulates the thickening of the lining of the uterus preparing the womb for a possible pregnancy. If the

egg is not fertilised, it dies and is shed, together with the lining of the uterus, ending the menstrual cycle. Table 5, summarises the phases of the menstrual cycle and the major role that different hormones play [111].

Phase of the Menstrual Cycle	Description of Physiological processes that occur	Hormones that determines the physiological changes
Menstruation Days 0-5	A new egg starts to grow in the ovary	**Follicle stimulating hormone** (FHS) is the prevalent hormone
Preovulation Days 5-13	The developing egg produces oestrogen	**Oestrogen** level peaks just before ovulation
Ovulation Days 13-15	The mature egg is released towards the uterus	A **luteinizing hormone** triggers ovulation
Postovulation	The endometrium is prepared for pregnancy	**Progesterone** levels reach their maximum

Table 5: Cycle phases of the menstrual cycle and hormones involved

As a woman approaches the menopause, the ovaries become less and less proficient and they eventually cease functioning. As a result, the woman's periods stop. The disturbance in hormone levels causes a number of symptoms and increases the risk of many diseases, as we describe in the next section.

5.2.2 Symptoms of the menopause, disease and remedies

Some of the most common symptoms of the menopause are:

Short term symptoms

Irregular periods: The ovaries become erratic in their production of oestrogen and progesterone, resulting in irregularities in the menstrual cycle and in the decline in fertility. This is usually the first sign of the menopause.

Hot flushes and night sweats: These can occur several times per day and often start with a feeling of increased pressure in the head that gradually spreads to the neck, then to the

shoulders and the chest. The result is disturbed sleep, headaches, muscle and joint pain, tiredness, lethargy and depression. This is the most common symptom of menopause, affecting about 75% of women [111].

Other symptoms: Sexual changes, as the oestrogen levels drop, result in dryness of the sexual organs, urinary symptoms, skin dryness and hair damage.

Long term consequences

Cardiovascular (CV) disease: Cardiovascular disease is the most common cause of death in the UK, being responsible for 100000 deaths (ischaemic together with cerebrovascular disease) in 1995 [112]. The risk of cardiovascular disease, for example atherosclerosis, is significantly higher for a woman after the menopause. In the States, it accounts for more than 50% of all deaths for women over 50. A possible reason is that the low density lipoprotein (LDL) cholesterol which encourages fat to accumulate in the walls of the arteries appears to increase in postmenopausal women [113].

Osteoporosis: This is a disease characterised by low bone mass and deterioration in the microarchitecture of bone tissue, that can lead to a significant increase in the risk of bone fracture. Osteoporosis, is prevalent in post-menopausal women, and the most common pathologies include vertebral, distal forearm and hip fractures [112].

Alzheimer's disease: It has been observed that this is more common in women than in men. It is suspected that there is a possible correlation between the disease and oestrogen deficiency.

From all of the mentioned above, one can understand the huge implications of the menopause for women's health, as well as the various health threats related to the induced hormonal imbalances. Many remedies have been proposed to relieve the symptoms related to the menopause and to minimise the risks of disease. These include exercise, healthy and careful nutrition, weight control as well as several recently suggested alternative treatments. However, to date, the most effective remedy for all the

symptoms and health threats related to the menopause is hormone replacement therapy (HRT).

5.2.3 Recommendation of HRT, benefits and risks

Hormone replacement therapy replaces the hormones that a woman's body ceases to produce after the menopause with natural hormones, which, unlike the synthetic hormones (used for example in oral contraceptives), have minimal effects on blood clotting. In clinical practice HRT is recommended:

- for menopause-related symptoms relief: HRT is effective in providing relief of climacteric symptoms (Hot flushes, sweats, vaginal dryness, headaches, migraines, depression, muscle pains etc.) and improving the general health of a woman;

- to prevent osteoporosis: HRT reduces the rate of bone loss and in some cases even increases bone mass. The extent of the improvement depends on the type of therapy and the route and dose used [111]. However, there have been studies which report no improvement in bone mass following HRT [114];

- another very important patient-group that needs HRT includes women who have had a hysterectomy. Over 60,000 hysterectomies (mostly aged between 40-49) are currently being performed in the United Kingdom each year [115], most of which aim to relieve women from heavy painful periods, chronic pelvic pain and pre-menstrual syndrome. A consequence of the operation is ovarian hormone deficiency, which often affects quality of life. Most of the symptoms of the menopause appear in women that have had a hysterectomy and are often more severe than in post-menopausal women. Although HRT is considered absolutely essential, at present a small number of hysterectomised women receive the therapy, the main reasons being the possible risks and the lack of a clinical protocol to regularly assess the hormone levels in these women.

In Table 6, we present a rough classification of "safe" against "not recommended" use of HRT depending on patient history, according to [112].

HRT considered "safe"	HRT "not recommended"
Patients with ischaemic heart disease, hypertension, diabetes (non-insulin-dependant)	Patients with breast cancer or breast cancer family history, endometrial cancer, severe liver or renal disease, history of venous thromboembolism

Table 6: "Safe" against "not recommended" HRT

There is some evidence that HRT can reduce the risk of heart disease; however this still has to be confirmed by large-scale, controlled clinical trials. There are at least three, ongoing, randomised, masked and placebo-controlled clinical trials related to the effects of HRT and the possible health implications:

- The women's Health Initiative (WHI) Trial is designed to uncover evidence about the effect of HRT on ischaemic heart disease, breast cancer, osteoporosis and Alzheimer's disease [116].

- The Heart and Oestrogen/Progestin Replacement Study (HERS) is designed to determine the effects of hormone therapy on fatal/nonfatal coronary events in 2763 post-menopausal women with documented heart disease who were followed for an average of 4.1 years [117, 118].

- The MRC trial is designed to assess the effect of HRT on the morbidity/mortality of CV disease, bone fractures, breast cancer and its impact on quality of life [119].

The most important aspects of HRT are increased risk of developing breast cancer, endometrial cancer and venous thromboembolism. However, the most common reason for stopping using HRT, are "less" serious side effects such as weight gain, headaches and breast tenderness. The lack of well controlled clinical trials and consistency in the planning of the experiments, as well as the small number of patients used in some studies,

are possible explanations for the inconsistent and sometimes controversial results that have been reported. In the next section, we examine in detail the effects of HRT in the breast.

5.3 Effects of HRT

The most important implications of HRT in mammography and the increased risk of breast cancer are:

5.3.1 Increased risk for breast cancer

There is a controversy in reports on whether HRT increases the risk of breast cancer. However, most studies agree that the risk of developing cancer increases for women that use HRT for more than 10 years. According to [120], the risk increases with the duration of use, while it is reduced after cessation. Similarly, the results from the Iowa Women's Health Study [121], suggest that there is a strong association between HRT use and breast cancer. A study concerning women that participated in the breast screening programme of West Scotland [122], reports that 12.3% of women with screening detected cancers and 22.2% of women with interval cancers were using HRT.

However, there is agreement that cancers due to HRT have a favourable prognosis, a fact that leads to the question: does this happen because different pathological processes occur, or because of increased clinical "awareness"? Holli et al. [123], compared a set of tumour aggressiveness indicators between HRT and non-HRT related cancers and the results suggested that breast cancer in HRT users is biologically less aggressive than those of non-users. Conversely, in [124], no correlation between biological factors and tumour prognosis was established, and the "favourable prognosis" was attributed to earlier detection through mammography.

On the other hand, there are reports (including [125]) that are sceptical about relating HRT to an increase in the number of mammographically undetectable lesions, while Speroff [126], after examining a number of reports, suggests (with some reservation) that any impact of HRT (in any form) on the risk of breast cancer is unlikely to be significant. Similarly, Lando et al. [127], did not find any significant statistical association between HRT use and the subsequent development of cancer.

Many studies have addressed and examined the possible risks of cancer (e.g. breast cancer, ischaemic heart disease and stroke, colorectal cancer, lung and ovarian cancer) related to HRT [128, 129]. We can conclude that although a consensus about the role of HRT in breast cancer has not been arrived at completely, most clinicians and researchers believe that women on HRT have an increasing risk of cancer. There is also agreement that the more time a woman uses HRT, the more risk she runs of developing cancer. The *Million Women Study* in the UK aims to determine (by the end of 2002), the relative risk of breast cancer between current HRT users and never users [130, 131].

5.3.2 Increase in density, changes in pattern

The response to HRT depends on the individual woman. As explained in section 5.5, the changes due to HRT are not necessarily homogeneous or global; in fact, they depend on the hormonal receptivity of the epithelial elements, thus there can be an increase in the density and a change in the density distribution (pattern) of the breast tissue.

Kaufman et al, suggests that women using HRT were found to have higher risk parenchymal patterns (according to Wolfe's criteria) than non-users [132]. Similarly, another study concludes that HRT users were more than twice as likely to have a high-risk pattern than never-users [133].

Several studies report that breast tissue density can increase in women who use HRT, sometimes causing cysts or fibroadenomas. Sterns and Zee [134] report that in a study of 1232 postmenopausal women, only 18% of the HRT group showed signs of continuing

breast tissue involution, as opposed to 38% in the non-users group where density decrease (involution) was apparent. Leung et al, found that 38% of the HRT users group had high breast densities, comparing to 11% in the non-users group [135]. In a hospital based study, Marugg et al, report a 31% increase in density for the HRT group, while only 14.3% of HRT users showed increased density in a screening-based study reported by the same authors [136]. Finally, in [137], the findings suggest that density changes due to HRT are dynamic and that: women who initiated HRT were more likely to show an increase in density, while women who stopped using HRT were more likely to show a decrease in density and women who continued HRT were more likely to show both increase in density and to *sustain* high density.

Other studies like [138, 139] also support the finding that tissue density can increase in HRT users. They classify such changes as focal, multifocal or diffuse mammographic increases in density. In addition, Cyrlak and Wong [140], report that oestrogen-only treatment promotes enlargement of cysts and fibroadenomas, combined treatment is more likely to lead to a diffuse increase in tissue density. Finally tissue density increase for HRT users is also reported in a number of studies including [141-143].

In a slightly different study, Lundstorm et al assess the density increase according to the type of therapy used [144]. They report an average 52% density increase in the group that received a combined therapy (oestrogen with progesterone), which was much higher than the 13% for the group that received oestrogen-only cyclic therapy and 18% for the group that received oestrogen-only treatment. Similarly, in another study it is emphasised that density increase in HRT users depends on the type of therapy used [145].

We may conclude that there is a consistent belief that HRT can increase breast tissue density. The type and degree of change depends on the exogenous hormone receptivity of the individual, and also on the type of hormones used in the treatment. An increase in tissue density can have an impact in screening mammography. This is discussed in the next section.

5.3.3 Decrease in mammographic sensitivity

Several studies have reported a decrease in mammographic sensitivity due to HRT. Kavanagh et al report a significant decrease in mammographic sensitivity and show that HRT users are more likely to have a false-negative report than non-users (odds ratio 1.60) [146]. Several other studies, including [135, 138], suggest that mammographic changes observed in some HRT users, can diminish the sensitivity of mammography for early breast cancer detection.

As a result of HRT increasing tissue density in postmenopausal women, the sensitivity of mammography can be reduced. That can have an impact on early diagnosis, as clinically occult cancers are more difficult to diagnose. As is discussed in the next section, many researchers address the possible dangers related to the use of HRT and suggest measures for monitoring users more carefully.

5.4 Recommended clinical management and the possible role of image analysis

The medical community, both in Europe and in the US, is concerned with the possible increase of cancer for women using HRT. In the UK, Sala et al, suggest careful clinical and mammographic follow-up for women on HRT, recognising the possible risks [133], while in the US, Berkowitz et al [147] suggest "serial" and "vigilant" monitoring of women that show a positive response to HRT (density increase). Other studies including [141, 143], strongly suggest more regular screening and the use of ultrasound in conjunction with mammography, for HRT users.

We conclude that there is global awareness of the possible risks associated with HRT. Although there are some disagreements about whether or not there is a statistically significant increase of the risk of breast cancer in women using HRT, all studies seem to agree that more frequent and careful screening should be introduced for women receiving treatment. These increased concerns about HRT users would inevitably increase the recall rate for incident screening, hence an increased cost and anxiety in the screening

population [122]. However, as reported by Harvey [148], a slight increase in the cost, due to increased use of diagnostic mammography and sonography, should not be a significant factor in the decision of whether or not to use HRT.

As is reported in [133], the mammographic patterns induced by HRT usage can be a new baseline, and changes with respect to this new pattern should be monitored over time. Once a woman responds to treatment and the breast density increases, temporal mammography becomes essential since an abnormality can be detected on the basis of significant difference with a baseline mammogram, as we described in the previous chapter. **Temporal mammogram registration and local tissue quantification can potentially provide the clinician with a *framework* for assessing the individual risk of cancer and maximising the chances for an early diagnosis of cancer.** This is the key idea underlying our work, as discussed in the next section.

5.5 Why the site of density changes is of particular interest

In the previous sections the issues relating to HRT were addressed mostly from an epidemiological point of view. In summary, once a woman responds to HRT with an increase in breast tissue density (to an extent that depends on her receptivity to exogenous hormones), the clinician typically suggests regular mammography and sonography in order to maximise the possibility of early detection of a cancer that might develop. In this section two approaches are taken to show that local density changes have a clinical importance: the first one is based on theoretical model for explaining tissue changes and the second on epidemiological findings (geographical distribution of breast cancers).

Figure 36, shows a simplified theoretical model illustrating the parameters that may induce tissue changes:

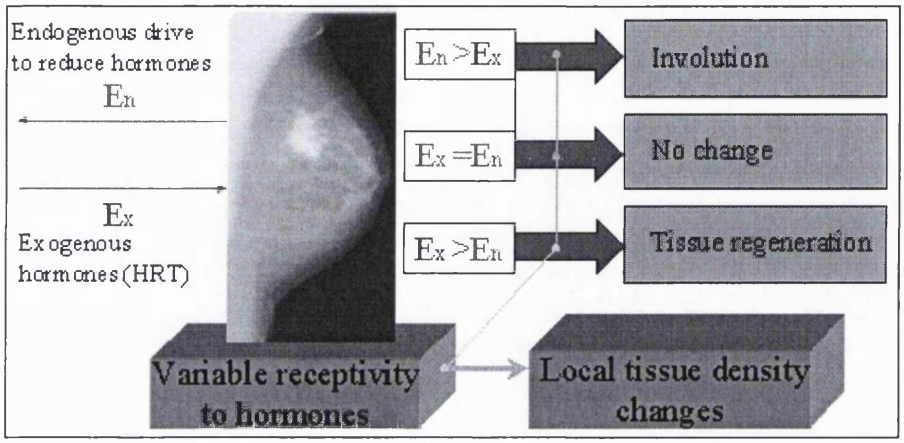


Figure 36: A simple model to explain breast tissue changes due to HRT

Oestrogen is a powerful mitogen (for both normal breast epithelium and breast cancer cells) and the induced mitogenic signal is mediated by the hormone receptors [163]. Since there is a variable receptivity to exogenous hormones it can be expected that the density increase should not necessarily have a uniform distribution. This can lead to a change in the mammographic pattern, as was mentioned previously. Local density increase (breast cell proliferation), may result in the development of a cancer. This leads us to consider that the location of a change in tissue due to HRT plays an important role when assessing the increased risk of cancer.

From a purely statistical point of view, certain sites in the breast seem to have different frequencies of cancer incidence: the upper outer quadrant (UOQ), of the breast is the most common site of cancers, accounts for 45% of detected cancers as is shown in Figure 37. The lower inner quadrant (LIQ) is the least frequently involved (~5% of cases) while the remaining locations of the breast (upper inner quadrant (UIQ), lower outer quadrant (LOQ) and central area) are intermediate in frequency [149]. In the same paper, blood flow studies in healthy women *showed an increased blood flow in the upper outer quadrant which was in accordance with the clinician's observation that* ***the UOQ seems to be the first region that shows a more active change from the postmenopausal fatty breast to an increase in density.*** The authors related those findings with a mammographically "perceived" local density increase in the upper outer quadrant.

In related studies, it is reported that 50% of synchronous and 37.5% of asynchronous bilateral breast cancers (in 19 breast cancer cases) were detected in the upper outer quadrant (UOQ) [150], while another study concerned with the location of primary neoplasia reported 60.7% of 146 cancers in the upper outer quadrant [151].

We conclude that local tissue density should be quantified and carefully monitored and examined particularly for women using HRT. We contend that a method for assessing local changes to HRT mammogram sequences could assist the clinician to assess the risk for cancer and contribute to early diagnosis. All of the above, converge to the belief that temporal mammography, could be the most effective early diagnosis framework for women using HRT. In the remaining of this chapter we explain our work on quantifying global and local tissue changes, while in the last chapter of the thesis, we propose some ideas for future work on local density quantification.

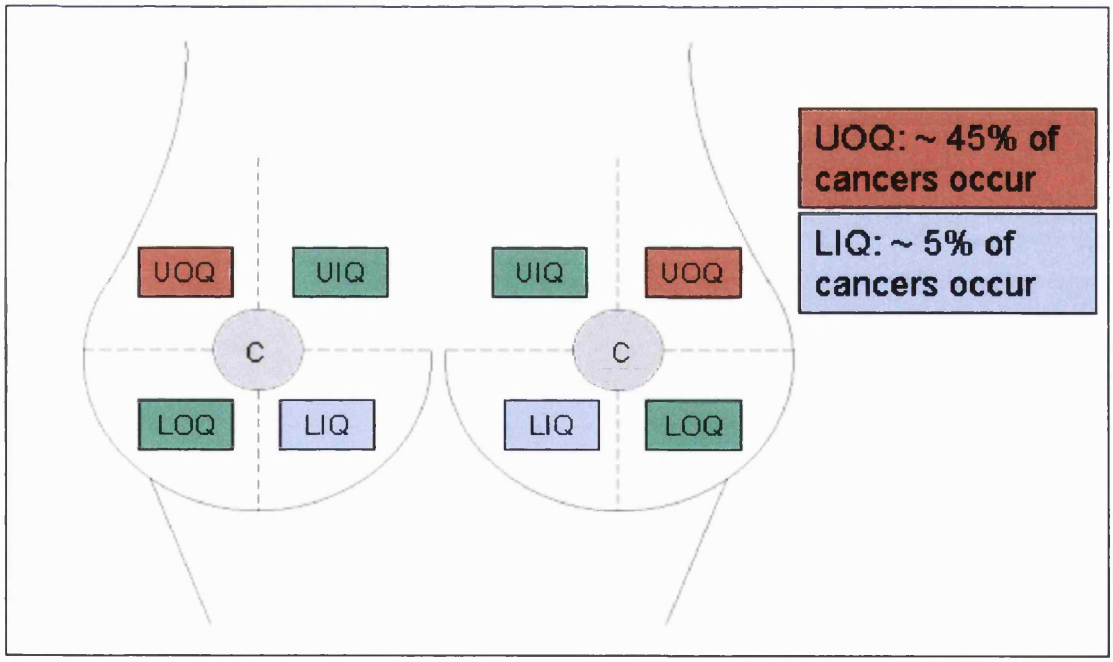


Figure 37: About 45 % of cancers occur in the upper outer quadrant of the breast (UOQ), while only 5% occur in the lower inner quadrant (LIQ). In the other quadrants (upper inner quadrant (UIQ), lower outer quadrant (LOQ) and central area(C)) the frequency in intermediate.

5.6 A proposed strategy, through a computer aided assessment of breast cancer risk, for women on HRT

The benefits of HRT cannot be doubted, yet the fear of developing breast cancer dissuades many women from taking the natural HRT hormones (oestrogen or combined hormones). Epidemiological studies provide valuable information about women taken as a whole. However, they these studies do not tell much about the likely incidence of disease, or the clinical management, for the individual woman. This is what the individual woman – and the clinician – is most concerned with.

To date, there is not a clinical protocol that combines the individual medical history of the patient and the help of computer tools to analyse temporal mammograms in order to assess whether the HRT patient is at increased risk of developing cancer. As noted above, several studies conclude that:

- Localised tissue changes (particularly breast cell proliferation in high-risk areas) may signal the development of a new cancer in HRT users. Image processing techniques such as registration, and quantification (for example the method presented in the next section) can assist the clinician to assess the local tissue changes consequent upon HRT use.

- If the mammographic tissue pattern changes due to HRT, it should become the new baseline mammogram. Again, image registration can assist the clinician to diagnose new growths on the basis of significant changes to previous mammograms. This can also help overcome the problems related to decreased mammographic sensitivity due to HRT changes.

We believe that temporal registration and quantitative comparison of HRT users mammogram sequences could provide the clinician with useful information to assess the risk of breast cancer for the individual woman. Figure 38, illustrates a possible patient

management protocol for HRT users. Global quantitative measures of density can assist the clinician in assessing the *global* response of the woman to therapy. If there is no density-increase the clinician can suggest another mammogram in one or two years, otherwise a more careful approach is needed: the "increased-density" mammogram should be considered as the new baseline, and the clinician should recommend more frequent mammography and ultrasound examination. Every time a new mammogram is taken, quantitative density measures can assist the clinician to understand if there is a "significant" increase in density while image registration can give him an idea about the distribution of the increased, "additional", density in the breast.

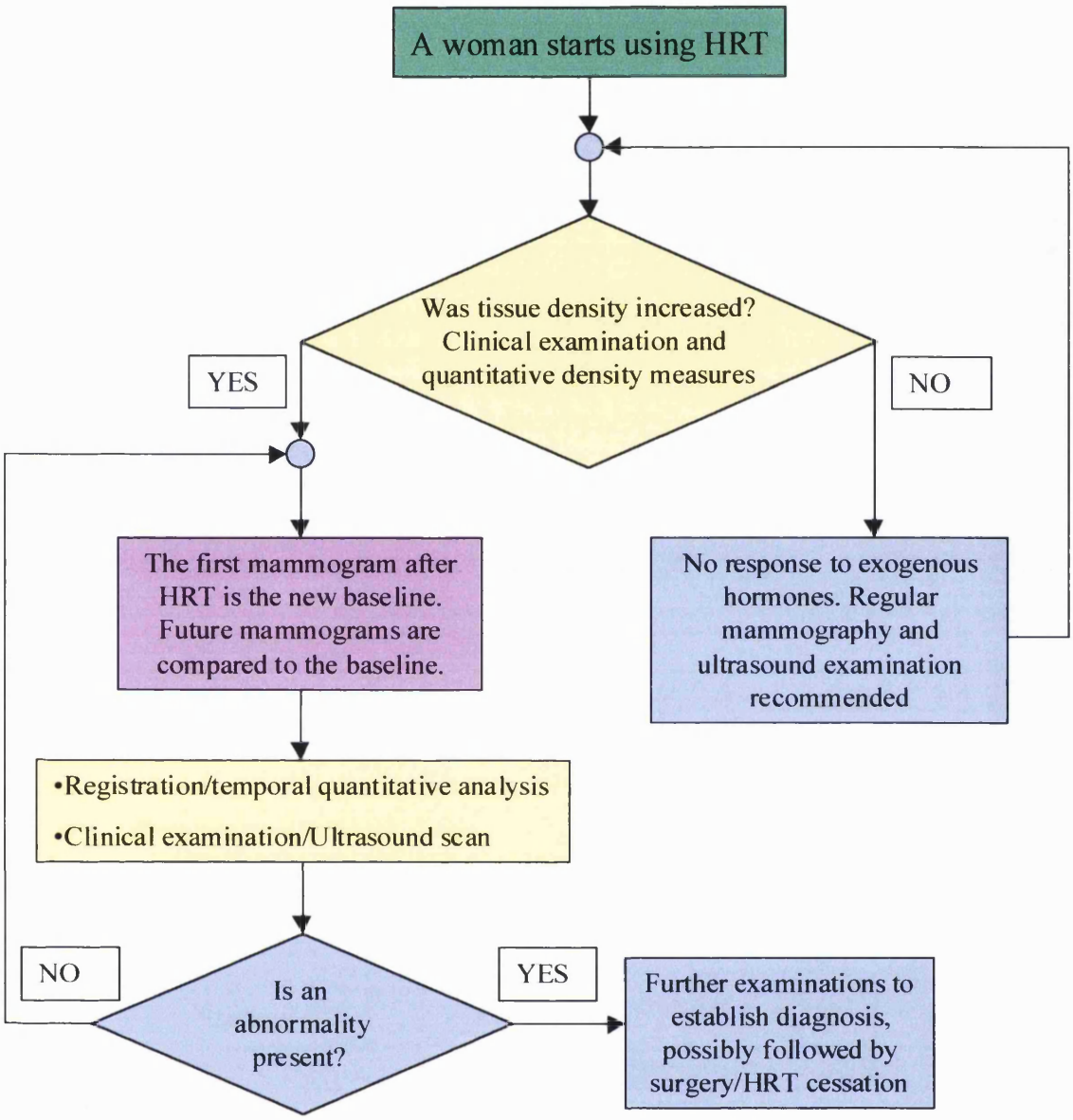


Figure 38: A proposed strategy for assessing the risk of cancer in HRT users. The elements of the flow charts coloured in yellow, are the ones where computer-aided mammography (temporal registration/quantitative comparison) could potentially provide additional clinical information. The feedback nodes represent the follow-up mammograms (which in the case of response to HRT should be more frequent than in the no-response case).

We need to consider the possible use of such methods (for example as part of the UK screening program) and evaluate retrospectively its accuracy in prompting the clinician in regions of potential abnormalities. In the next section, we present our novel method for accurate quantification of tissue change in temporal mammogram pairs.

5.7 A novel method for robustly quantifying mammographic changes in temporal HRT sequences

5.7.1 Previous work

Describing and measuring HRT induced tissue changes is a part of almost every clinical study that examines data obtained before and during therapy. The clinical aim is to correlate such changes with an increased risk of cancer or to the type of therapy used, and to assess the possible influence on diagnosis. The latter consideration, means that increased glandular density induced by the therapy can decrease mammographic sensitivity [152], since the observed local "brightening" of the film can make diagnosis of lesions more difficult.

The most usual change observed on a mammogram is regeneration or increased density in fibroglandular tissue but it is not uncommon to observe the opposite or to observe no change at all! A simplified mechanism explaining the variability in the observed changes would consist of the indigenous drive to reduce hormone production competing with the induced (exogenous) hormones combined with the different hormonal uptake between individuals.

To the best of our knowledge, there is no previous computer analysis-assisted framework that combines non-rigid registration (alignment) of temporal mammograms with accurate breast tissue density quantification. In addition, previous efforts to quantify breast tissue density and correlate it with the risk of breast cancer were based on poor breast density measures [153-155]. Such measures effectively calculate the area of fibroglandular (fibrous and stromal tissue) in the mammogram or, in the best case, the percentage area of the mammogram covered with dense (fibroglandular) tissue. That is to date the

predominant method used for tissue quantification (both in Europe and in North America [153-156]). We consider such a method inadequate for a temporal quantitative comparison of HRT mammograms for two reasons:

1. It only considers the area and not the intensity of fibroglandular tissue in the mammogram. The fact that the denser a tissue segment of the breast is, the higher the x-ray attenuation is not taken into consideration. For that reason, density area (or percentage area) can be inaccurate since it only measures the spatial extent of the projected tissue.

2. For temporal studies, we have to take into consideration the difference in compression between the two acquisitions. As has been reported in [4], breast structures move and deform differently under different compression and according to properties such as tissue density and elasticity. For this reason, area density measures of the same woman can change with different compression, even if there is no actual change in density.

5.7.2 Introduction to our proposed tissue quantification method

We have introduced a novel method for assessing mammographic (density) changes based on temporal analysis of mammograms of women using HRT. A two-step approach is taken for quantification of density change in HRT image sequences:

- First, we normalise the intensities using the h_{int} representation [4] (discussed in the next section and in Appendix C). This allows us to estimate the amount of non-fatty tissue in the column of the compressed breast above each pixel. Without registering the images we can assess *global* changes in non-fatty tissue between two successive scans. Global change experiments are discussed in section 5.8.

- In order to assess local, rather than global changes, we need to align the images in the same co-ordinate frame. We use the registration method presented in the previous

chapter to pre-process (align) the mammogram sequences. The idea of quantifying local changes is discussed in section 5.9.

5.7.3 Extracting robust tissue density quantitative measures from mammograms

Our aim is to extract quantitative tissue-density information directly from the mammograms in order to assess global or local changes due to HRT. Because of the relatively weak control over the image acquisition process, it is difficult to eliminate variability in image characteristics, such as contrast and brightness. The differences in imaging conditions lead to a non-rigid transformation between the *intensities* of temporal pairs of mammograms. For this reason, it becomes difficult to compare temporal mammograms, and in particular to relate the temporal intensity differences of dense regions to the actual breast-tissue changes due to HRT.

The h_{int} representation of interesting tissue [4], was briefly introduced in chapters 3 and 4. Its main advantage is that it "normalises" mammogram pairs since each pixel value represents the amount of 'interesting' (non-fat) tissue between the X-ray source and the image. A short description of the h_{int} representation can be found in Appendix C. Figure 39, is a schematic representation of the calculation of h_{int}. This representation provides an objective framework for quantitative comparison of dense breast-tissue.

Changes of fatty to glandular tissue are precisely changes in non-fatty – i.e. h_{int} – tissue. Hence, h_{int} is the natural representation to use for the quantitative approach needed. By using the h_{int} representation as a first step in our analysis, we can construct measures of "interesting" tissue change that reflect the underlying tissue changes due to HRT. This is tested in the next section, where we design an experiment to evaluate the accuracy of the proposed quantitative density measures.

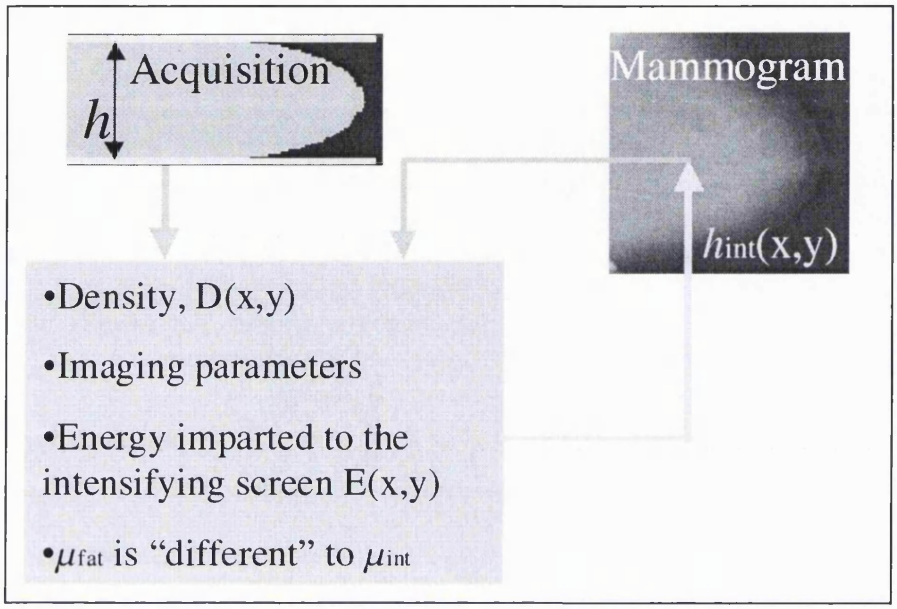


Figure 39: A representation of the calculation of the h_{int} representation. Mammogram information (pixel density values) is combined with acquisition information and parameters to calculate the "height" of interesting tissue in each pixel.

5.8 Design of a global tissue-quantification validation experiment

5.8.1 Introduction

It is difficult to correlate quantitative measures of dense tissue to the changes in the breast over time. Our specific interest in this chapter is to assess changes in tissue over time in women on HRT. As mentioned in section 5.7.1, measurements of the area of dense tissue in the mammogram have been used to correlate tissue density with the risk of developing cancer. *Our goal has been to investigate the relationship between the clinician's perception of global changes in temporal HRT mammogram pairs, and the calculated measures.* For these specific experiments the images were not registered since we wanted to test the "agreement" between the clinician and our suggested quantitative measures, on a *global* basis.

Similar to large-scale clinical studies as in [157], an experienced clinician described the changes in each HRT temporal mammogram pair. The HRT sequences were classified in three categories by the clinician as shown in the following table:

Assessment of change according to the clinician	Involution (less glandular tissue observed in the most recent mammogram)	Regeneration (more glandular tissue observed in the most recent mammogram)	No change observed between the temporal HRT PAIR
Number of mammogram pairs in each category	***21***	***21***	***17***

Table 7: Data used for quantification experiments

We used this table as the "ground truth" for comparison with our quantitative measures. We note some weaknesses of the experiment:

- "Ground truth" was based on the opinion of only one clinician
- The clinician compared the images directly from the computer screen instead of the light-box that he was used to
- Some of the changes were described as "slight" or "very slight" (e.g. very slight involution) and the distinction between those changes and the "no-change" categories was not obvious to the clinician.

For these reasons, it was expected that the quantitative measures calculated for the 17 mammogram pairs of the "no change" category would include values found in other categories (in other words a value corresponding to a very slight change could be found in the "no change" category since the distinction is not clear). This reflects the difference in describing a change as insignificant between the clinical expert and the quantitative measure; the clinician examines the temporal sequences visually and assesses the overall changes as insignificant, but in reality there is always a change in the breast-tissue.

For the reasons mentioned above, we first examine only those cases where there is either involution or tissue regeneration, and the results are presented in the first part of section 5.8.3. We also examine the cases that the clinician described as "no change", in the second part of the same section.

Registration is not necessary for assessing global changes in tissue composition, rather for the local detection and description of changes as is discussed in section 5.9.

5.8.2 Measures used

In the experiment, we used three measures of tissue density, one being the projected area of glandular tissue, the other two being based on the h_{int} quantitative analysis to calculate differences over time from HRT mammogram sequences:

1) Projected area

As mentioned in section 5.7.1, the projected area of glandular tissue has previously been used as a measure of tissue density. Figure 40 illustrates the "conventional" method for isolating the area of glandular tissue, called interactive thresholding. In this method, the user has to define a global threshold that includes as much "density" as possible. The percentage of density in the whole mammogram is used to describe the density content of the breast. Although, the area of glandular tissue is a measure claimed to be correlated with the clinician's perception of tissue density, it does not include intensity information (and even if it did it would not be consistent since that depends on the imaging parameters).

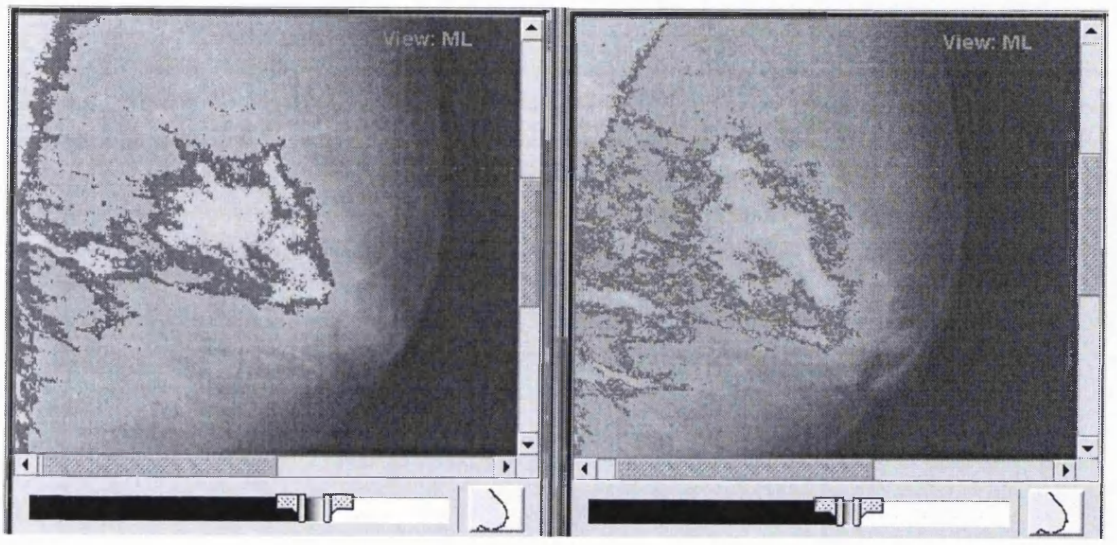


Figure 40: Interactive thresholding is performed in order to include all the dense tissue in each mammogram.

As a mammographic measure, the area of the projected glandular tissue suffers from two basic flaws:

- It does not account for the thickness of the projected tissue since the resulting intensity is not directly proportional to the thickness and can differ significantly due to the imaging conditions.

- The interactive segmentation of the "dense tissue" is a subjective procedure and therefore the measure obtained is not consistent.

However, the nature of this measure is closer to the clinician's perception of changes since the segmentation of "dense" tissue and the description of changes in two mammogram sequences are similar processes.

2) Sum and Volume of interesting tissue

After normalising the HRT sequences using the h_{int} representation of interesting tissue and segmenting the pectoral muscle, we compute the following measure to describe global changes in the composition of the breast:

a) The sum of "interesting" tissue:

$$\Delta h_{int}\ (\%) = \frac{S_1(h_{int}) - S_2(h_{int})}{S_1(h_{int})} \cdot 100, \text{ where } S_i(h_{int}) = \sum_x \sum_y h_{int}(x_i, y_i) \text{ for i=1,2}$$

b) The volume of "interesting" tissue:

$$\Delta V_{int}\ (\%) = \left(V_1(h_{int}) - V_2(h_{int})\right) \cdot 100, \text{ where } V_i(h_{int}) = \frac{\sum_x \sum_y h_{int}(x_i, y_i)}{H_i \times A_i} \text{ for i=1,2}$$

Where S_1, S_2 are the sum of interesting tissue values in each image, A_1, A_2 the area of the images (number of pixels), H_1, H_2 are the heights of the compressed breast in each mammogram and Δh_{int}, ΔV_{int} our measures of tissue change. It is important to include the same part of the breast when calculating S_1 and S_2. In addition, the first measure is arithmetically more stable to calculate than the second which is divided by the height of the breast under compression, H, (whose calculation or estimation can be prone to an error).

5.8.3 Experiments and results

First experiment: Involution and tissue regeneration

By using this description as the gold standard for characterising change between the HRT sequences the rule shown in Table 8 was used to evaluate our method for a total of 49 HRT sequences.

Change in interesting tissue	Corresponding change in breast tissue-density due to HRT
$\%\Delta h_{int}$ or ΔV_{int} or $\%\Delta$ Area >0	***Density Decrease*** Fibroglandular to fatty
$\%\Delta h_{int}$ or ΔV_{int} or $\%\Delta$ Area <0	***Density Increase*** Fatty to glandular/stromal

Table 8: The rule to assess mammographic changes due to HRT, from measuring the % normalised difference in "interesting" tissue between the HRT sequences.

Even using a rule as simple as that described in Table 6 we have achieved good agreement between our measures and the expert's opinion. Figures 41 and 42 present some of the results in graphs, where the degree of agreement between the clinician and our measures can be visualised.

For temporal mammograms showing tissue regeneration of the fibroglandular tissue we achieved a 71% agreement with both the Δh_{int} measure and the ΔArea (Figures 41(a) and (c)). In those cases, the ΔV_{int} measure did not perform well (34% agreement), probably because it describes the percentage of interesting (non-fatty) volume in the breast and can be influenced by other changes in the breast (e.g. water content, changes in total volume) and because this measure depends on the compression height H, which is difficult to accurately calculate.

For temporal mammograms showing involution of the fibroglandular tissue, we achieved an 81% agreement using the Δh_{int} measure and 86% using the ΔV_{int} measure (Figures 42(b) and (c)) and 57% using the area difference ΔArea.

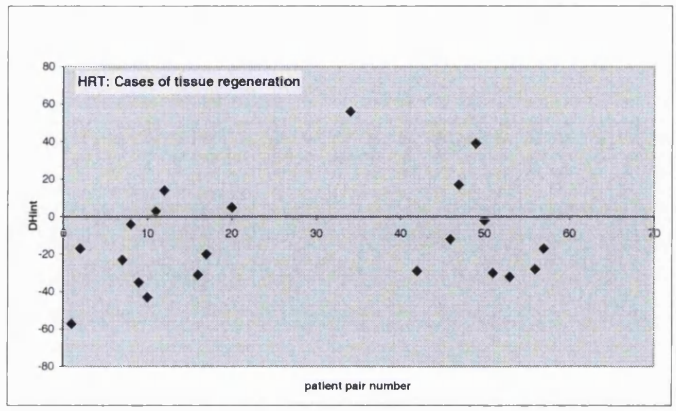
Table 9, summarises these results. Overall, the measure that performed the best is the Δh_{int} measure, as it gave consistently good results (agreement with clinical "ground truth") in both cases (involution, tissue regeneration pairs). This experiment confirms,

that h_{int} is the right representation to use for the quantitative comparison of temporal mammograms.

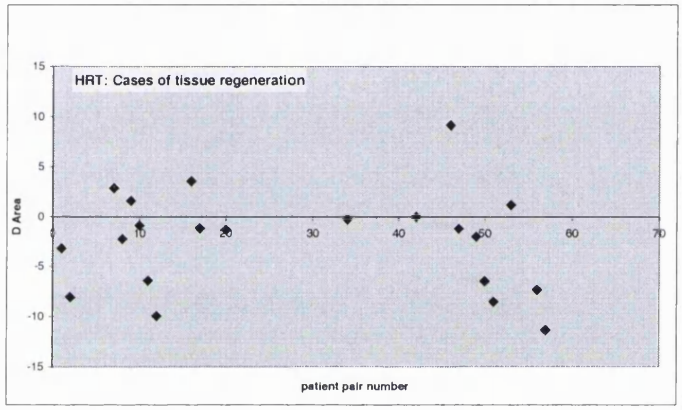
It also has to be noticed that the clinician examined the original (not intensity-normalised), mammograms. It would be useful for future experiments to use both intensity-normalised temporal mammograms (h_{int}) and registered sequences.

Density change measure	$\%\Delta h_{int}$	ΔV_{int}	$\%\Delta$ Area
Agreement with clinician in "tissue regeneration" pairs	71%	34%	71%
Agreement with clinician in "involution" pairs	81%	86%	57%

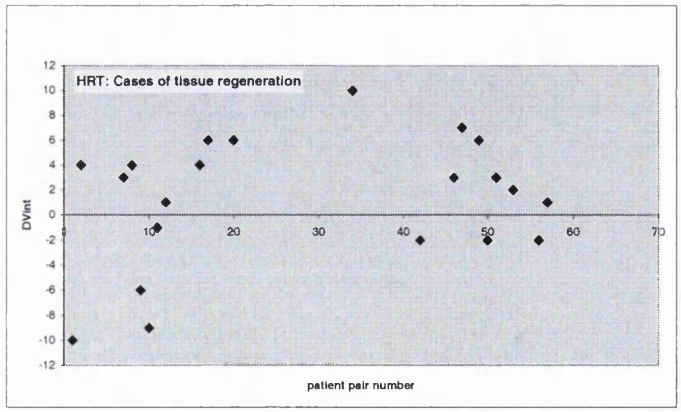
Table 9: Results on the agreement between the quantitative measures used and the clinician, concerning tissue changes



(a)

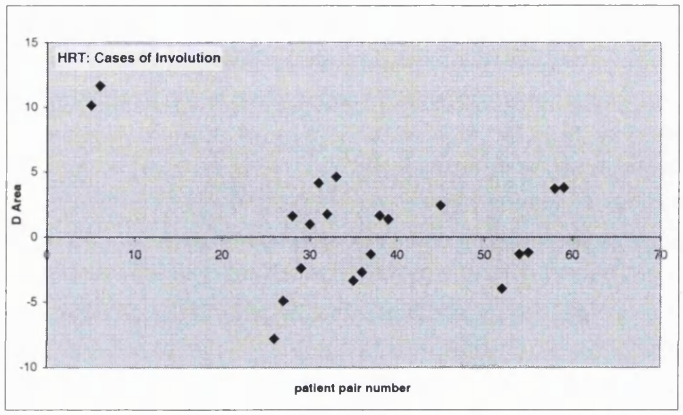


(b)

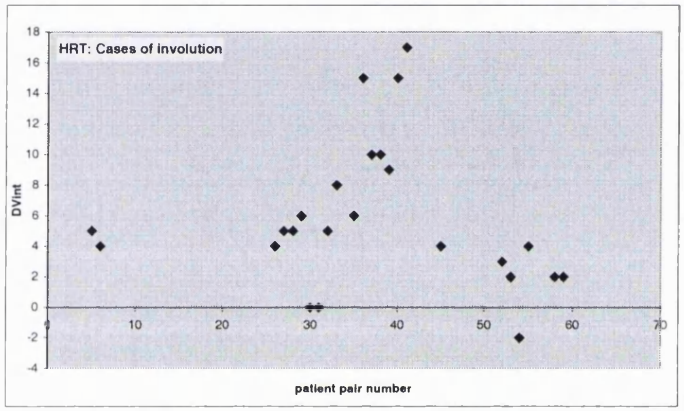


(c)

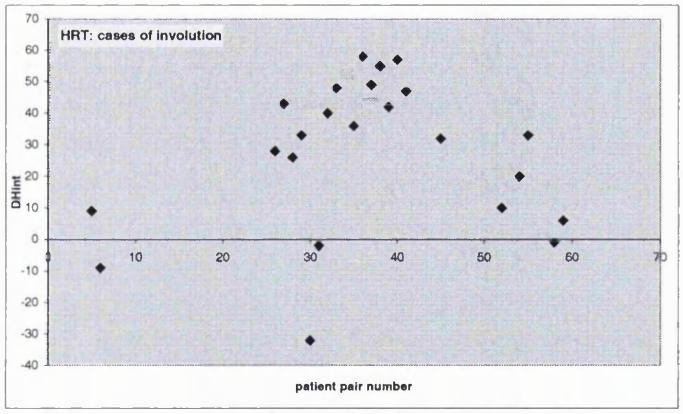
Figure 41: The cases described as "tissue regeneration". The measures of change are *expected to be negative* in order to match the expert's description (according to Table 6). (a) Shows the (%) ΔArea change measurements, (b) shows the change in the volume of "interesting tissue" ΔV_{int} measurements and (c) the change (%), in the "sum of interesting tissue" Δh_{int}. ΔArea and Δh_{int} were the most successful measures both resulting in an overall 71% agreement with the clinician's description (according to the rule in Table 6).



(a)



(b)



(c)

Figure 42: The cases described as "involution". The measures of change are *expected to be positive* in order to match the expert's description (according to Table 6). (a) Shows the (%) ΔArea change measurements of glandular tissue while, (b) shows the change in the volume of "interesting tissue" ΔV_{int} measurements and (c) the change (%) in the "sum of interesting tissue" Δh_{int}. ΔV_{int} and Δh_{int} were the most successful measures showing respectively an overall 86% and 81% agreement with the clinician's description (according to the rule in Table 6).

Second experiment: What happens when "no-change" is reported?

As mentioned in section 5.8.1, the clinician was asked to classify all the pairs in three categories, the third being those cases for which "no-change" was observed. However, in reality:

- there is always a change in the breast
- the clinician described changes (involution or regeneration) as definite, "slight", and "very slight"

In theory we would expect differences in quantitative measurements to be close to zero. However, this is not the case especially since the clinician's assessment between "very slight change", and "no-change" is not guaranteed to be consistent. The quantitative results for the "no change" pairs are presented in Figure 43. It is interesting to note that there is a zone of change in each graph that includes most of the measured changes for the pairs that were described as "no change" by the clinician. (In Figure 43 (b), it is interesting to observe that 70% of the mammogram pairs are constrained in the ±3.5% range of change of volume of interesting tissue). We can assume that there should be a zone of uncertainty around zero, which quantitatively will include cases of small tissue changes. More extensive experimentation is needed to estimate the "ranges" of quantitative values that correspond to very small tissue content change. Nevertheless, we need to mention that it is very difficult to correlate objective (computer measures) to subjective (clinician) descriptions of the "no-change" cases. One obvious obstacle is that human perception finds it hard to be consistent in the characterisation of very subtle mammographic changes over time (e.g. very slight involution vs. no change). *However, clinically, there is no particular interest in extending experiments of "no change" since there is no tissue change of underlying clinical importance (while HRT induced density increase can be "suspicious").*

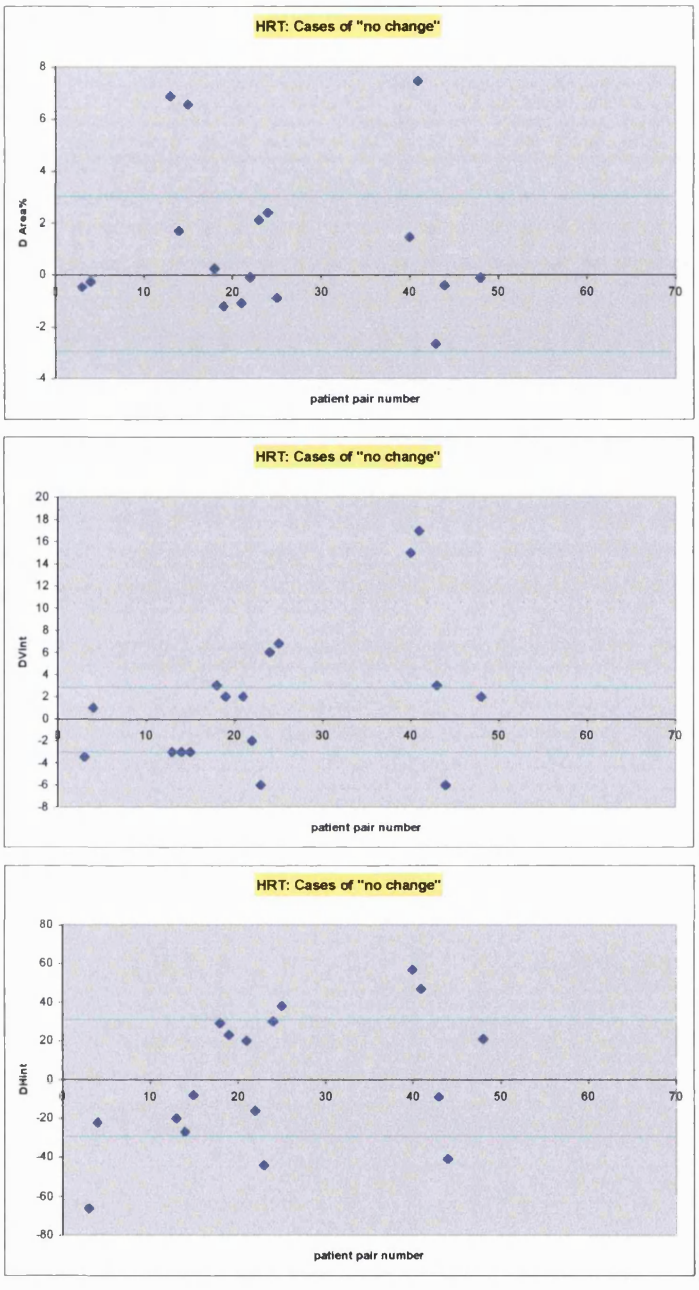


Figure 43: The cases described as "no-change". The measure change is expected to be near zero. The blue lines suggest an uncertainty region in each case, which reflects the fact that there is always a slight change. (a) Shows the (%) change ΔArea measurements of glandular tissue while, (b) shows the change in the volume of "interesting tissue" ΔV_{int} measurements and (c) the (%) change in the "sum of interesting tissue" Δh_{int}.

5.9 Can we quantify local changes using registration?

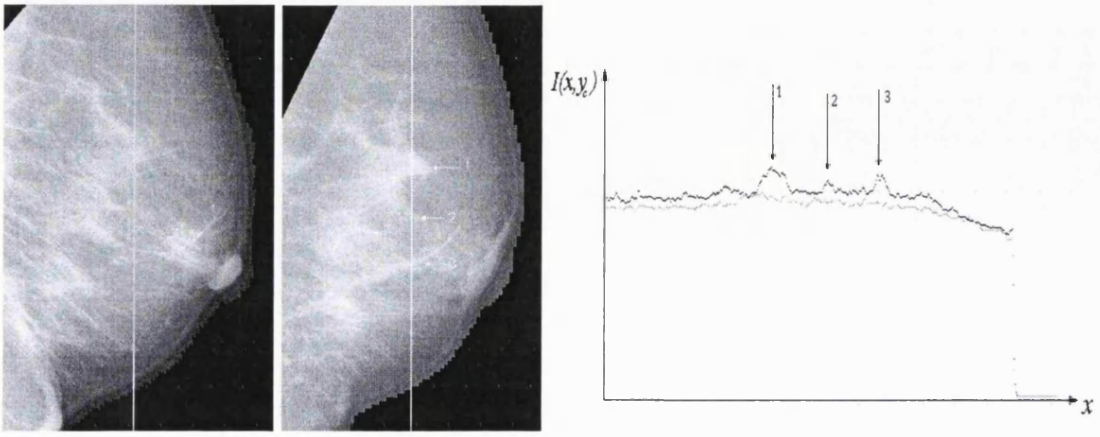
The quantitative description and comparison of changes in mammograms of HRT users, and our eventual objective of detecting and describing such changes, is not only destined to assist the clinician to visualise change, but to help to understand the aetiology of these changes as well. For example, one of the breasts that is starting to develop cancer, may possibly exhibit a different (and sometimes even the opposite!) response to therapy with respect to the other. Similarly, a localised increase in density (e.g. in the upper outer quadrant) may be a sign of a cancer developed due to hormonal stimulation.

In the previous section we suggested measures of global changes in tissue composition. Our ultimate goal is to be able to assess significant *local* changes in tissue density. This could be done by combining registration and quantification in the same framework, so that the clinician can quantitatively compare temporal or bilateral HRT mammograms locally. Figure 44 (a) and (b), shows a registered and normalised mammogram sequence of a woman on HRT. It is obvious in the second image that there is a "global" increase in glandular tissue. Using the Δh_{int} measure the change was calculated to be 35% increase in glandular tissue. The line-intensity profile *after* registration in Figure 44 (c) shows that there is indeed an increase in the overall intensity with some "larger" local variations (numbered 1,2,3) that correspond to the three areas of glandular tissue regeneration indicated in figure 44(b). *More registration results on HRT sequences are presented in the Gallery of Results (Appendix E).*

However, in order to assess local changes we need to transform one of the images to the co-ordinate system of the other, and still be able to use image based quantitative measures of tissue (based on the h_{int} representation in our case). That sequence of operations is not trivial, and there is an additional issue introduced. The transformation applied to align the images can "shrink" or "extend" the mammogram image, depending on the difference in compression without the intensity being corrected. One can think of this problem in the same way as a deformation of an elastic material constrained in a horizontal plane. If pressure is applied the material is deformed but in such a way that the

volume is preserved. This does not happen when we transform 2D images using standard methods (like Radial Basis Functions), since there is no change in the intensities and the sum of intensities across the image changes after applying the transformation.

Having this in mind, one can think of the mammogram as a 3D structure (the third dimension being the h_{int} quantitative value, or, given standard imaging parameters, the pixel intensities represented in grey-scale). When we want to register it with another acquisition of the same patient that was taken with different compression applied to the breast, we need to transform (including scaling) the mammogram in a non-rigid way (with the method described in chapter 4). This way, the "volume" of the 3D structure (mammogram) is not preserved and although we have aligned the images, any image based-quantification (like the h_{int} representation) analysis is prone to error. We have developed a method to overcome this by approximating the appropriate intensity transformation so that the "volume" of the mammogram is preserved after applying a spatial transformation. This is presented in the next section and completes the necessary image-processing infrastructure proposed for local quantification and comparison in HRT sequences.



(a) (b) (c)

Figure 44: Registration and intensity comparison of an HRT mammogram pair. (a), (b): An HRT mammogram pair where in (b) density is increased due to HRT. *Using the Δh_{int} measure the global change was calculated to be 35% increase in glandular tissue.* The images have been normalised (using the in h_{int} representation) and registered, (c): The line-intensity profiles showing the intensity distribution along the same bright line indicated in (a) and (b). The darker one corresponds to (b) and comes in agreement with the observed increase in glandular tissue in image (b). "Larger" local variations (numbered 1,2,3) correspond to the 3 areas of glandular tissue regeneration marked in (b).

5.10 Volume preserving elastic transformation for local breast-tissue quantification

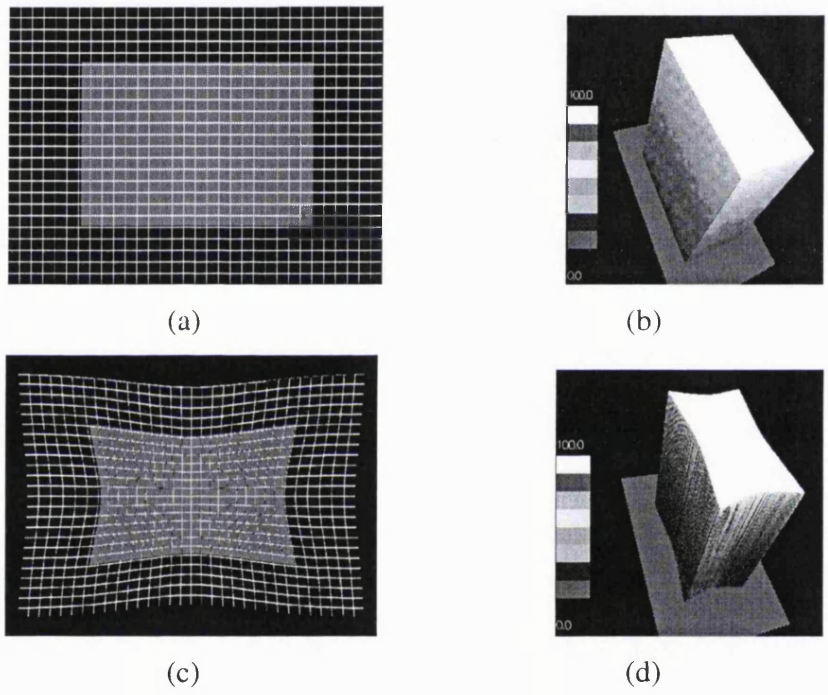
5.10.1 Introduction

It has been demonstrated that in order to quantitatively compare a pair of mammograms, there is a need to register the images. However, the elastic registration process involves pixel rearrangement and scaling which can significantly alter the total "volume" of the mammogram image.

The latter can be defined as: $$V_{\text{Image}} = \sum_{i=0}^{N} I(x_i, y_i)$$

Where N is the number of pixels across the image.

This problem is illustrated in Figure 45, where the "volume" of a synthetic image is reduced to 70% of the original, after applying a non-rigid transformation.



(a) (b)

(c) (d)

Figure 45: (a): The original rectangle undeformed, (b): the original rectangle plotted as a 3D surface, (c): Image (a) is deformed using thin-plate spline interpolation, (d): The intensity remains constant and as a result the total image volume is reduced by 70%.

In many cases, due mainly to temporal differences in the breast size/compression, the mammogram size and geometry can change significantly between acquisitions. Consequently, aligning temporal data can reduce (or increase) the image volume. This is not important for "un-normalised" mammograms since the intensities of corresponding regions are not necessarily related. However, in the h_{int} representation, every pixel value represents the height of "non-fatty" tissue at that point during acquisition. In this section

we demonstrate an intensity correction method based on the registration deformation field that compensates for the volume error introduced by elastic deformation.

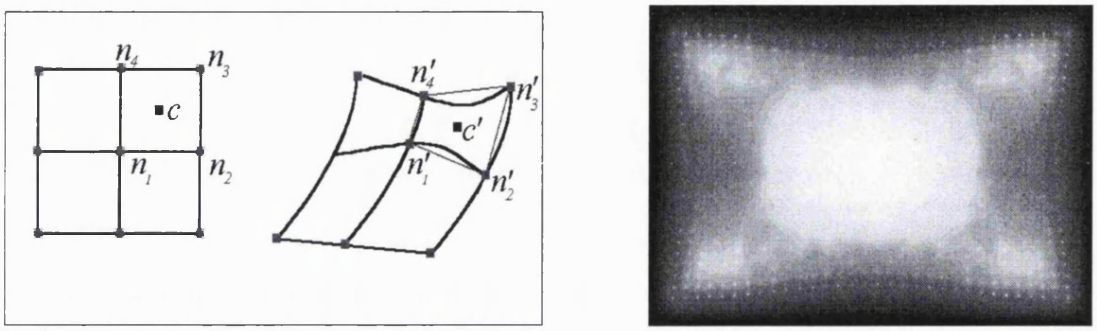
5.10.2 The "volume-preserving" transformation method

A 2D elastic transformation produces an undesirable effect in h_{int} images where the "height" of each pixel, and consequently the image volume, are meaningful quantitative measures of the breast anatomy. Our method for intensity correction can be summarised in the following idea: if we take a square cell in the image to be transformed and calculate a "deformation measure" for that cell, then we can "adjust" the intensity (height) of the deformed cell so that the "volume" is preserved. In Figure 46 (a), we illustrate this concept. The cell represents the image before registration. The central node c, and the 4 surrounding nodes n_i, (i= 1, 2, 3, 4) are transformed to c' and n_i' according to the image based calculated transform.

Based on those points, we can calculate the area of the cell before and after the transformation (we use a local affine approximation of the transformed region). The ratio of the areas is then used to assign a modification value to the transformed node c':

$$M(c') = \frac{A}{T(A)}$$

Where A is the area of the orthogonal cell, T is the applied transform. After we calculate the modification value for all the nodes of the image, we calculate an intensity displacement field using finite differences (Laplacian interpolation). This is shown in Figure 46 (b). In order to compensate for the image "volume" changes, we multiply the image intensities by the intensity correction field.



(a) (b)

Figure 46: (a): A local affine approximation is used to describe the deformation of an orthogonal image cell (defined by the nodes n_1, n_2, n_3, n_4), (b): The calculated intensity-correction field based in all the nodes of Figure 45 (a) and 1 (c).

The remaining volume (error volume) can be equally distributed along the image pixels so a 100% volume preservation can be achieved. This requires the calculation of an adjustment ("plateau") value according to the formula:

$$p = \frac{V_{original} - \sum_{i=1}^{N} I(x_i, y_i) \cdot C(x_i, y_i)}{N}$$

Where p is the plateau value, $V_{original}$ is the volume before the transformation, I is the image and C the intensity-correction field. In the next section we illustrate this concept with a synthetic example followed by a real patient case.

5.10.3 Results and conclusion

In Figure 47 (a), the same rectangle as Figure 45 (b) (of constant intensity) is deformed using thin-plate spline interpolation. However, the intensity of the square does not change, resulting in loss of volume as it is shown in Figure 47 (b). Figure 47 (c) is the corrected image using the method described in the previous section. The maximum intensity rises from 100 to 173 (on an 8-bit grey-scale) and the intensities are distributed

according to the intensity-correction field shown in Figure 46 (b). The error volume is distributed equally in all pixels and the resulting volume is the same as in the unreformed image.

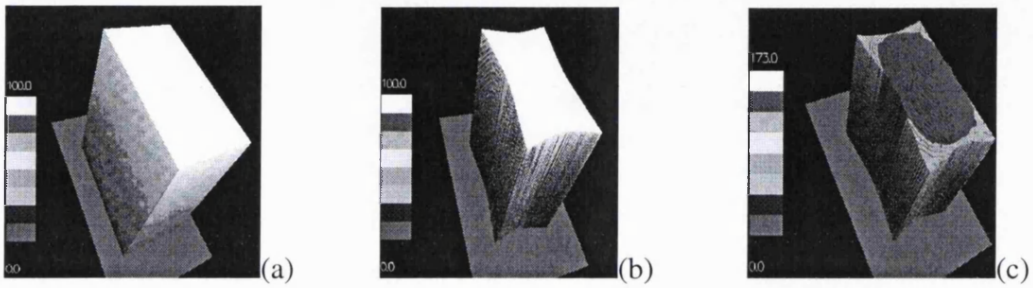


Figure 47: The intensity correction algorithm for the deformed square in Figure 45. (a): The original square, (b): After transformation the volume is reduced by 30%, (c): Using the intensity correction field (Figure 46 (b)) and the "plateau" value p, we correct the intensities so that the image volume is preserved.

In Figure 48 (a) and (b), we show a temporal HRT pair. The difference in breast compression (and size) between the two acquisitions is significant. When the images are registered, the image "volume" is reduced by 33 %. By using the mammogram registration method the images are aligned and then the intensity correction fields are calculated, first using only boundary points, and then using internal landmarks to better approximate the deformation (Figures 48 (c) and (d)). The corrected images for these two registration scenarios are shown in Figure 49 (b) and (d) respectively.

In conclusion, our method aims to preserve the h_{int}-based anatomical information content after registration, so that local quantitative comparison is possible. As shown in Figure 49, the result of the intensity correction depends on the complexity of the calculated registration transformation. For this reason, in Figure 49 (d) the glandular tissue is more pronounced as the internal glandular areas are taken into consideration in the registration process. In contrast, Figure 49 (b) calculates a smoother transform (and intensity correction) based only on the boundary alignment. The accuracy of the corrected mammograms depends on the accuracy in the calculation of the geometrical

transformation that relates the temporal mammogram pair. The corrected image is calculated mainly for accurate comparison of tissue density in temporal HRT pairs. However, it still needs to be clinically meaningful and in our future work we aim to validate that assumption. An appropriate experiment to validate the intensity preserving alignment method presented here, is to use differential compression data [4] (the patient is kept in the same position, and the breast is imaged at the same time, in two different compression levels). This could provide a ground truth of the intensity correction field, since (unlike temporal mammograms where the breast changes between acquisitions) the resulting h_{int} images (after registration and intensity correction) should be exactly the same. In the next section we describe a preliminary validation experiment using differential compression mammography data.

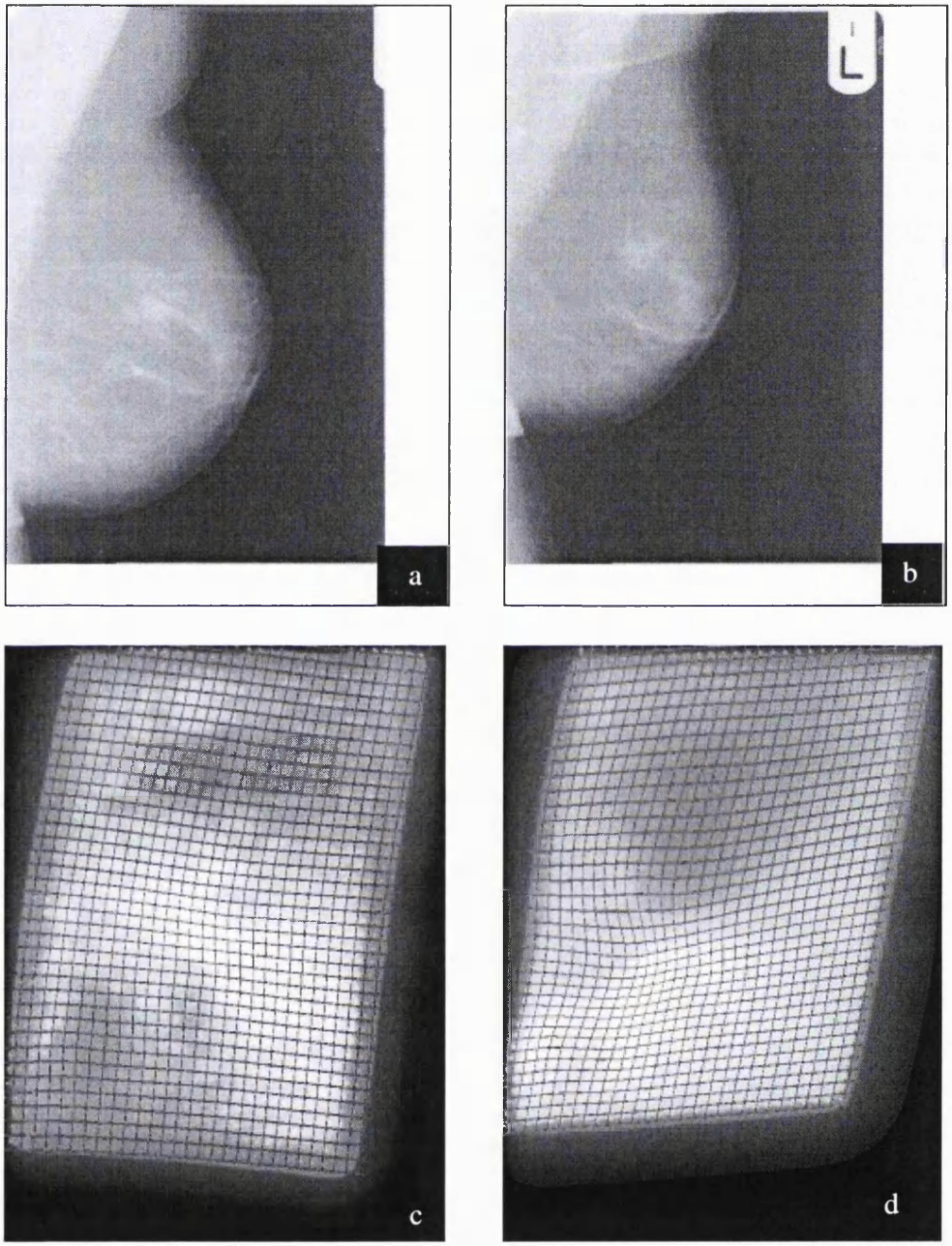


Figure 48: (a), (b): A temporal HRT pair, where the breast compression is significantly different, (c), (d): The intensity correction fields for boundary-based, and internal landmark registration respectively

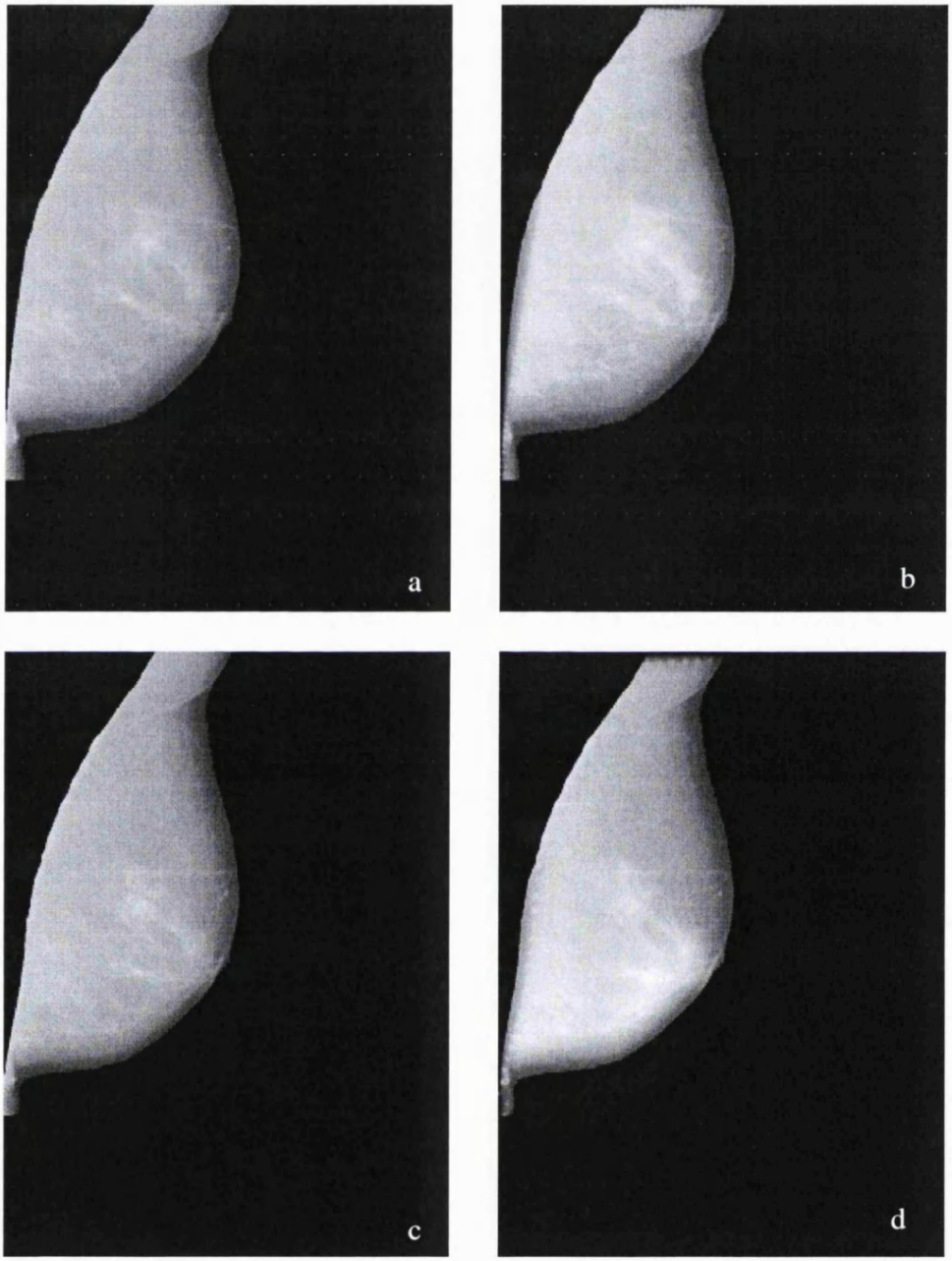


Figure 49: (a): The mammogram in (Figure 4 (a)) registered to the one in Figure 4 (b) using only boundary points, (b): Using that transformation we calculate the intensity corrected image, (c): The same as (a) but using internal landmarks to better approximate the deformation, (d): Again based on the calculated transformation we calculate the intensity corrected image.

5.10.3 A validation experiment

Summarising what was mentioned in the previous section, to locally compare HRT sequences we need to represent the mammogram pair with an anatomically significant model (i.e. the h_{int} representation) and register the images including the intensity correction previously described. However, it is very difficult to validate the accuracy of this technique since the actual differences in a mammogram pair can't be known *a priori.* For this reason, we intend to use differential compression mammogram pairs, where the image intensity distribution and geometry are expected to be identical after the application of our techniques. A first experiment is described in this section.

Figures 50 (a) and (b) illustrate a differential compression mammogram pair where the tumour (near the nipple) is further compressed in (b), resulting in a larger spatial expansion and a lower average intensity. In Figure 50 (c), the second mammogram (Figure 50 (b)) has been registered to the first one (50 (a)), while in Figure 50 (d), the registered image is corrected using the technique described in the previous section. The difference images in Figure 51 show a significant overlap (in Figure 51 (b) after registration but without intensity correction), which is maximised in the difference image shown in Figure 51 (c) (after registration and intensity correction). This is because the bright region near the breast edge and the intensity differences in the region of the cancer almost disappear after the intensity correction (zero is represented with the grey value of the background which is rescaled according to the differences in the subtracted images).

In addition to calculating the difference images, we calculate the image volume in the region of cancer after registration and intensity correction. The cancer region is manually segmented from the mammograms shown in Figures 50 (a) (less compressed mammogram), 50 (c) (second mammogram registered to 50 (a)), and 50 (d) (intensity corrected mammogram 50 (c)). The segmentations before and after intensity corrections are exactly the same. The image "volume" quantitative results for the whole mammogram and the segmented region of interest (cancer) are summarised in Table 8. It is noticeable, that the volume difference of the segmented cancer (the manual

segmentation was constant before and after correction) is reduced to 4% after correction (from 19% before correction).

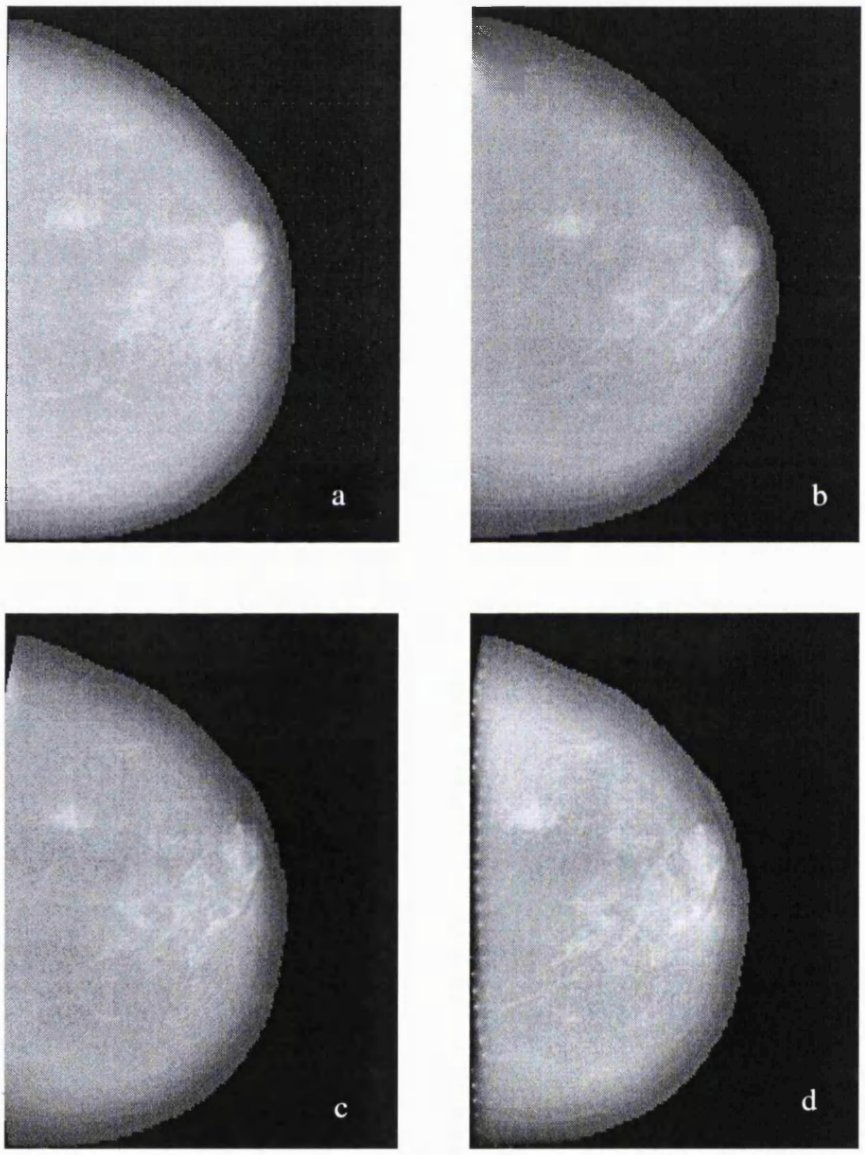
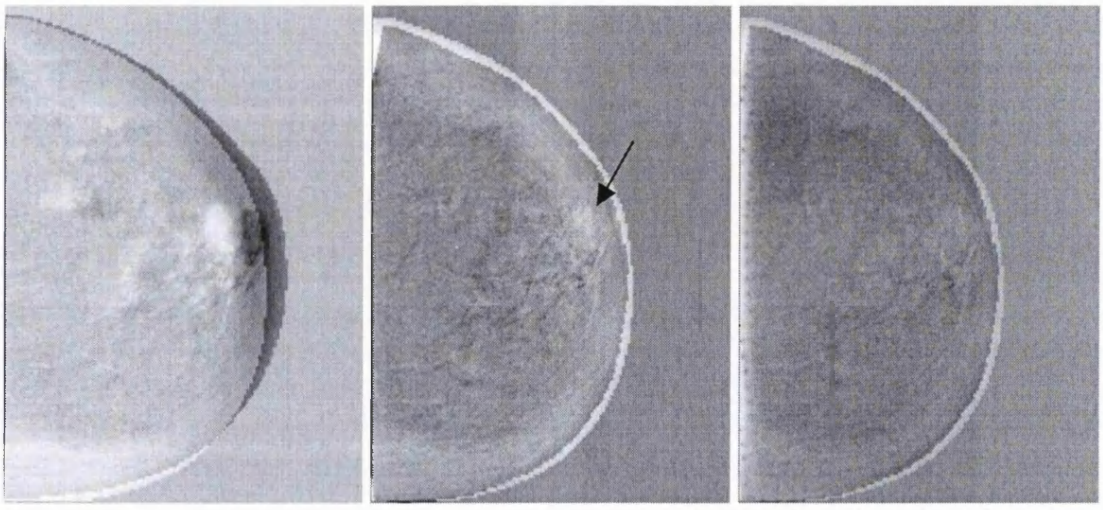


Figure 50: (a), (b): A differential compression mammogram pair, represented using the h_{int} representation, (c): Image (b) aligned to image (a), (d): The registered image (c), after applying the intensity correction. Images (a) and (d) are expected to be identical.



(a) (b) (c)

Figure 51: Difference images in the differential compression pair (Figure 15 (a) and (b), (a): Before registration, (b) after registration and (c): after registration and intensity correction. Note that after intensity correction the difference image is more homogenous and the intensity misregistration in the area of the cancer (pointed with an arrow in image (b)) almost disappears.

Volume units in Pixels x (8-bit greyscale)	Target Volume (Figure 15 (a))	Volume before correction (Figure 15 (c))	Volume after correction (Figure 15 (d))
Whole mammogram	29439392	24856420	29251438
Segmented cancer	716235	581858	685431

Table 10: Quantitative comparison of the whole mammogram volume and the segmented cancer before and after intensity correction. The units are in grey-levels after normalisation. First we sum all normalised intensities for the whole mammogram and then for the segmented region of interest (cancer).

5.11 Limitations and alternative methods for evaluation

In this section, we summarise the basic limitations concerning the design and implementation of experiments for evaluating the proposed methods for global and local quantitative measures of density. Some of these problems have already been mentioned in the previous sections. However, here we discuss all the limitations in evaluating quantitative mammographic techniques and suggest some alternative methods that could provide a more objective validation of any technique for quantifying mammographic density.

Firstly, the presented experiments in automatic classification of global tissue have some intrinsic limitations. As mentioned in section 5.8 some basic limitations include the fact that only one clinician classified the changes and the viewing of the mammograms that was done in a conventional computer screen (instead of the lightbox). Moreover, we must address the problem of establishing a reliable ground truth. Although, the experiment presented in this thesis is a simple one, we believe that a more careful design is necessary in order to ensure the objectiveness and consistency of the perceptual descriptions of density. This is further discussed later on this section.

The image processing framework presented in sections 5.9 and 5.10 for local quantification of density is intrinsically difficult to validate. Unlike the case of global density description it is not trivial to perform a similar experiment on a local basis. Especially because the tissue content of the breast constantly changes, which means we can't predict the actual changes in anatomical content over time. This would require very accurate segmentation of similar regions in both mammograms and subsequently a similar effort (as in the global case) for establishing ground truth of density change via clinical description. Unfortunately, this procedure would be prone to interpretation errors.

Fortunately, in this thesis we were able to use a mammogram pair of differential compression data (where two mammograms are taken in the same session with slightly different compression). This means that we had two mammograms with exactly the same

tissue composition and this served as ground truth, since the local quantitative analysis was expected to give identical results in both mammograms. As was shown in section 5.10, the results were satisfactory. However, the fact that differential compression data are not widely available (due to the increased radiation dose) poses a limitation for using such data for validating local quantification techniques.

For all these reasons, we need to suggest alternative methods for testing either the proposed or future techniques for global or local mammographic density quantification. First of all, in order to design a more robust experiment we will have to investigate:

- The inter and intra observer variability when assessing density changes. This would require the clinician to examine the same set of mammogram pairs several times until it can be established with some certainty that his/hers description of changes is *consistent*. In addition, it would be useful to investigate the variation in describing density changes between two or more clinicians.

- The consistency in describing density changes when the clinician examines the mammogram films on the "lightbox" and the digital mammograms in the computer screen.

A more robust way for establishing ground truth in the description of density changes could be possible if instead of density changes in a pair the clinician would characterise the density in each mammogram individually. This is because there are some standard ways in classifying mammogram patterns and many experienced radiologists are quite familiar with them. This would make the "ground truth" more robust but the inter and intra-observer variability as well as the variability between film screen and computer screen assessment, should be investigated in this case as well. Some of the most popular techniques of describing mammogram patterns is the Wolfe classification [193] and the recently established BI RADS (Breast Imaging Reporting and Data System) from the American College of Radiology [194]. Both techniques take into consideration both the

density and the distribution of the density in the mammogram and could be very useful to evaluate any mammogram quantification framework in the future.

5.12. Summary and conclusions

For women who decide, or are advised, to take HRT, it is important to monitor the response to therapy so that the increased risk for developing cancer is assessed. The reported density increase in HRT users has given rise to concerns for increased risk for cancer [158, 159]. This risk depends on whether or not the exogenous hormones stimulate glandular tissue regeneration in certain locations in the breast. In this chapter, we explored the idea of quantifying local breast-tissue density changes and comparing mammogram density over time. We presented results from several clinical studies that show that although still controversial there are many concerns about the use of HRT. In particular, postmenopausal women who use HRT for more than five years run an increased risk of developing cancer due to the "regeneration" of fibroglandular tissue that is often induced by the exogenous hormones. We presented a method that combines mammogram normalisation and volume-preserving registration that can be the starting point for temporal-local breast tissue quantification.

First we introduced a novel method for assessing mammographic (density) changes in HRT mammogram sequences. The temporal sequence of mammograms is represented using the h_{int} representation of interesting tissue. The % change in h_{int} (non-fatty tissue) in the pair is a good approximation of the fibroglandular change in the composition of the breast tissue. The expert's description of the changes in the HRT sequences of our data set was used as the gold standard for evaluating our quantitative measures. The % change in h_{int} was found to be the most reliable measure for describing global changes in density.

Our aim is to be able to *locally* describe the induced changes in the breast so that regions that exhibit strong glandular regeneration could be highlighted and the clinician would examine them more carefully in order to early diagnose a potential abnormality. In order to assess these changes locally, we suggest the combination of quantitative analysis with

temporal mammogram registration. Though, image registration can significantly change the image "volume" and therefore compromise the accuracy of the desired local quantitative comparison. To overcome this problem, a volume preserving registration has been developed, which preserves the anatomical content of the h_{int}-represented mammograms after alignment. We presented results on both synthetic images and temporal HRT mammogram pair. The method was also evaluated in a differential compression mammogram pair.

We believe that quantitative measures of local tissue change can assist the clinician in determining the increased risk of cancer for a woman who uses HRT. More experimentation and validation work is needed to prove that our suggested framework for temporal analysis of HRT images can play that role.

5.13 Final statement about the presented work

In this chapter we presented a method for quantifying density changes in HRT mammogram sequences. Our results are very encouraging since there is a good agreement between the suggested measures and the clinical perception. Although in our experiments the h_{int} representation of interesting tissue performed better than the other measures further experiments are required to prove the hypothesis that this representation can *accurately* describe the *actual* tissue changes in the breast, over time. Most of the limitations and suggestions of more sophisticated experiments were discussed in section 5.11.

Chapter 6: Discussion and future work

6.1 Conclusions

The work presented in this thesis has focused in two directions: mammogram registration and quantitative comparison of HRT sequences. Tables 10 and 11, summarise the aims of our work, the motivation behind them and the most significant conclusions:

Aim	To investigate the requirements and implement a method for reliably registering mammogram sequences.
Motivation	Temporal comparison of mammograms is a common practice in the reading centers of the screening programme. Registration as a part of a computer aided mammogram interpretation system is expected to improve the breast-cancer detection rate of the screening programme and help to equalise standards throughout the UK.
Conclusions	1. The breast boundary is the most reliable feature for mammogram alignment. 2. Three maximum curvature points are calculated in the breast outline for each mammogram. These points most often have an anatomical significance (corresponding to the axilla, nipple and rib points), and allow us to establish correspondence *consistently* in any pair of mammograms. 3. The boundary-based registration is a good first approximation (especially in large deformations). Using a multi-scale segmentation algorithm a small number of regions can be detected and matched. The final registration is done using an approximation scheme. 4. Our validation experiments show that there is a significant improvement in the correspondence of features in the mammogram pair after registration. In addition, the boundary landmark detection rate is satisfactory and can be improved in many ways. 5. By aligning temporal mammogram data it is easier to understand the evolution of regions of tissue over time (e.g. notice that a "suspicious" dense region corresponds to scar tissue due to previous surgery). In addition the difference image after registration can reveal areas of significant change. 6. Using registration, interval cancers can be studied retrospectively by examining if there were signs of the cancer in the corresponding location (given by the registration transformation) in previous mammograms.

Table 9: Aims, motivation and conclusions for the registration work presented in this thesis.

Aim	To apply quantitative analysis tools (h_{int} representation) and registration to temporal mammograms of women on HRT.
Motivation	Quantification of global changes provides information about the woman's response to exogenous hormones while local tissue changes could assist the clinician in assessing the risk of cancer due to hormone-induced breast-cell proliferation.
Conclusions	1. HRT can induce tissue changes that change the mammographic patterns of the breast. In such cases, the new pattern can be used as a baseline for comparison to subsequent mammograms using registration. 2. Based on the h_{int} representation measures for quantitative comparison of breast tissue can be derived. The accuracy of these measures in describing HRT changes was validated using the clinical assessment as the ground truth. The results showed a good correspondence of these measures in assessing *global* changes in breast tissue. 3. Quantitative assessment of *local* tissue changes can have an important clinical impact since it would help the clinician to better understand the underlying changes due to HRT and assess the risk for cancer. 4. Combining the h_{int} representation with mammogram registration is a first step towards local tissue quantification. However, especially if the mammogram size significantly differs between the two mammograms (e.g. because of changes in size or compression), the quantitative image "volume" is altered due to sub (or over) sampling. 5. In order to preserve the quantitative information after registration, an intensity correction method based on the deformation field, was introduced. A validation experiment on differential compression data (the *only* ground truth of "equal volume" mammogram sequence), showed satisfactory results. 6. The suggested work is a basis for a clinical system for temporal global and local tissue change assessment.

Table 10: Aims, motivation and conclusions of the quantitative comparison work presented in this thesis.

The research presented in this thesis has been inspired by clinical problems related to breast cancer. Our clinical colleagues strongly believe that advances in medical imaging can provide them with crucial information for breast-cancer diagnosis. However, to date most of the modalities for breast cancer imaging provide complementary information leading to the need for data fusion tools. Mammogram registration is a basic step towards a future clinical system capable of multi-modality registration for breast cancer imaging.

6.2 Limitations

The most general limitation for mammogram image analysis is the large variability of shape and mammographic appearance glandular tissue structures. In addition, poor control in imaging conditions result in non-rigid intensity transformations between temporal mammograms. The most important limitations of the work in this thesis can be summarised in the following:

- The breast-edge segmentation presented in the second chapter is a very fast approximation based on the histogram of the image. However, in some cases the intensity distribution in not clearly distinguished from background noise resulting in losing part of the edge after thresholding. Although a sophisticated breast segmentation was not of main concern in our work we believe that a more accurate segmentation can improve the robustness of our proposed registration framework. Nevertheless, the trade off between speed and accuracy should be taken into consideration, especially because image analysis should be fast in a digital reading environment where the clinician needs to examine a large number of patients (around 100) per hour.

- There is no doubt that inclusion of internal landmarks improves the accuracy of registration. However, so far we have used only the center of mass of the matched regions. Though this is a fast way to include them in the registration it does not fully exploit the shape and orientation information of these regions. To be able to match exactly each individual region would be a complex problem and again, the speed of the registration could be affected without a corresponding increase in accuracy.

- Our novel work on HRT sequences has shown encouraging results and because of the clinical importance of the role of the hormones in breast cancer, any advance in the HRT image analysis can have an impact in the clinical assessment of the increased risk due to hormones. However, the main limitation is the inability of mammography

to provide information related to the type of proliferated cells (e.g. *atypical vs. normal cells).* Nevertheless, information about local density increase would alert the clinician and the woman will be further examined (histopathology on a fine needle aspiration biopsy sample).

6.3 Ideas for future work

All suggested ideas for improvement should be considered together with performance requirements for the related application (e.g. In a digital reading system, time is at a premium since the radiologist has to examine a large number of mammograms per hour). Most of the improvements suggested in this section aim at the development of a clinically useful system. However, in this final chapter we only discuss the image analysis issues and not all the other technical issues related to the development of a system that can be used in clinical practice.

The main ideas we suggest for future work are:

- Improvement of the breast-outline segmentation: This can improve the detection rate of the axilla, nipple and rib points. There are two approaches that can improve the segmentation: the first is to try and separate the breast outline from the background on a local basis (rather than defining a global threshold like we the method we used) and the second is to try and enhance the region near the breast edge where the contrast (with respect to the background) is usually very poor.

- Improvement of the anatomical significance of the boundary landmarks: The curvature detection algorithm is dependent on the sampling rate of boundary points used to calculate the spline (coarser sampling preserves the detail of the segmentation result, while coarser smoothes it). In chapter 4, we proposed a method to overcome this problem by examining the curvature maxima in different sampling resolutions. These landmarks often have an anatomical significance which can be improved if further constraints (to the maximum curvature ones) are added in the detection

algorithm. Alternatively (or in conjunction), an improved segmentation of the boundary could improve the detection rate as well.

- A specific algorithm for nipple detection: Unlike the rib and axilla points, there is additional information to be exploited for a more robust calculation of the nipple. In addition, in some cases the nipple does not appear at all at the breast outline. Future work on nipple detection could take into consideration the fact that glandular structures (ducts) converge to the nipple (e.g. through segmentation of glandular tissue) in order to be able to detect the location of the nipple in mammograms where it does not appear near the breast outline.

- Additional work on HRT sequences: There is no doubt that computer-assisted assessment of tissue density changes has the potential to impact on the clinical decision making related to the use of HRT for the *individual* woman. In section 6.6 some ideas concerning future work on HRT mammograms are discussed.

- Data fusion: The complementary nature and information content of medical images from different modalities dictates the development of *fusion* algorithms able to register multi-modality breast images. Two data fusion problems related to mammography are discussed in section 6.5.

- Tissue classification using texture: One very interesting image analysis problem is to automatically classify the different types of tissue in the breast. This can provide additional information/constraints for our proposed work on registration and tissue quantification. Our preliminary work on tissue classification is discussed in the next section.

6.4 Initial results on texture segmentation

6.4.1 Introduction

The aim of this work is to segment each mammograms in a small number of distinct classes with anatomical significance (e.g. a "rough" classification of the image in density fat and the pectoral muscle). To date, it has not been possible to automatically classify all the different types of tissue directly from the mammogram. An automatic segmentation of different mammographic tissue-regions can be for the following reasons:

- Segmentation of the pectoral muscle: By using a method based on texture, we do not have to make any assumptions about the shape of the muscle (e.g. linear approximation of pectoral muscle) since the area exhibits a highly homogenous and high intensity texture.

- Breast boundary segmentation for registration of temporal pairs: The background has consistently different second order statistics from the breast image and the average intensity is significantly lower. The only region that can be difficult to locate is the fatty region near the breast edge. Again, enhancement of this region (as in [107]) could help overcome the problem of "overlap" between the breast edge and the background.

- Automatic segmentation of glandular tissue: This, in conjunction with the h_{int} representation, can further improve our quantitative temporal HRT image analysis since we will be able to compare only the regions of the image that correspond to the projected glandular tissue.

- Automatic segmentation of the radiological tag: This could be based on the homogenous texture of that image feature and on the fact that is distant to other segmented "classes" (e.g. fat and density).

6.4.2 Method

First the image is divided in small patches (15x15 pixels). In each patch normalised second order statistics are calculated, for example the correlation measure:

$$f_3 = \frac{\sum_i \sum_j ijp(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

where $p(i, j) = P(i, j) / R$ is the normalised joint probability of the pixels i and j. For each image patch i, a texture vector $\underline{T_i}$ is calculated and all the vectors are classified in a desired number of classes using hierarchical clustering. Although ideally there should be three classes (pectoral muscle, fat and density), in practice there are outliers that can appear as a distinct class. This problem poses a limitation for the use of textural clustering. However, we believe that these kind of classification problems can be solved with the inclusion of basic heuristics.

6.4.3 Results

In Figure 52, the pectoral muscle is segmented by classifying all the patches in two categories. Figure 52 (a) shows the original mammogram, while in (b) we see the clustering of the patches corresponding to the pectoral muscle, against the rest of the image. There is a danger that some patches inside the breast (in this case they would have to be homogenous texture and high intensity, as the pectoral muscle) can appear as a distinct class. As mentioned before, this could be corrected by including simple heuristics.

In Figure 53, a temporal pair (Figures 53(a) and (b)) is shown. The corresponding textural classification is shown in Figure 54 ((a) and (b) respectively). Again, in Figure 54 (b) part of the pectoral muscle is classified in the same category as glandular tissue but this can be corrected by adjusting the number of classes or introducing some basic heuristics (e.g. including the radial distance from the origin in each pattern will assist in segmenting only the pectoral muscle).

6.4.4 Discussion

The method can be very useful as a first tool to decompose complex mammographic images, but further investigation is required together with the inclusion of heuristics in order to deal with outliers. The results appear to be very promising for the segmentation of the pectoral muscle and the areas of glandular tissue as well as for the radiological tag. However, the breast edge is not accurately segmented (the method detects the glandular disc rather than the real breast edge) since the area close to the edge is not detected.

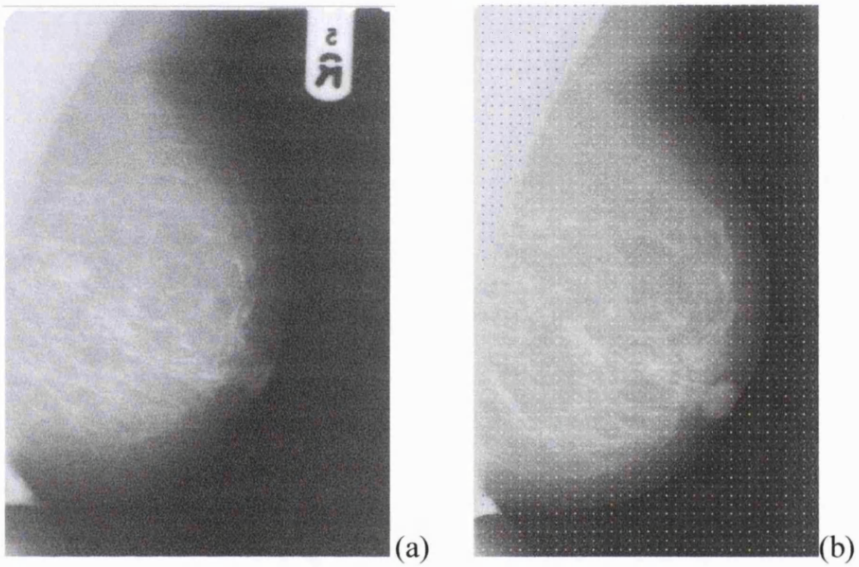
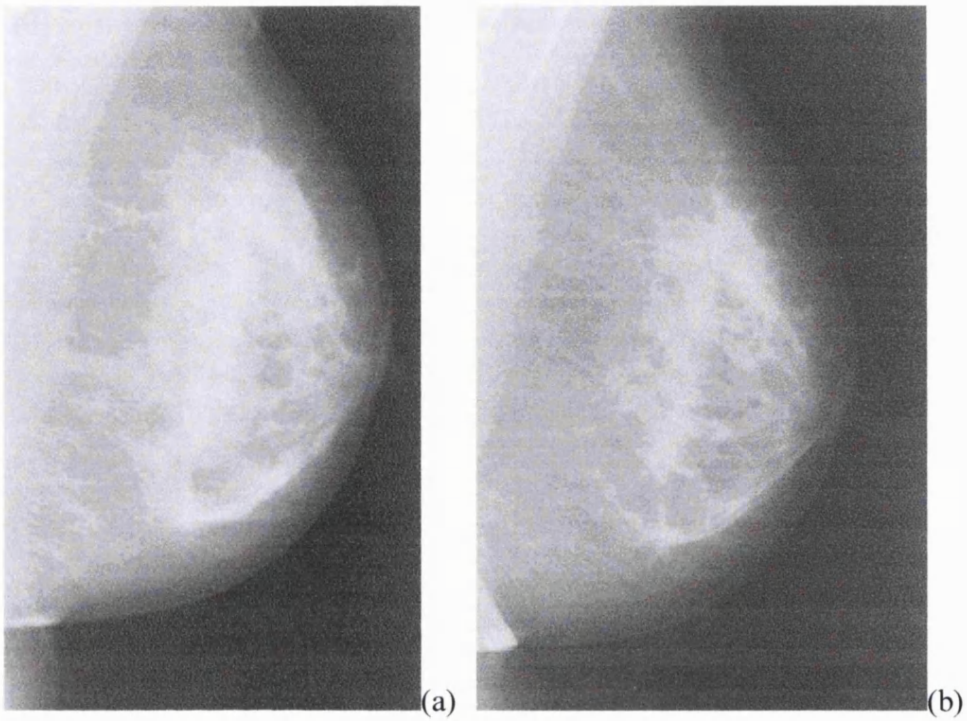
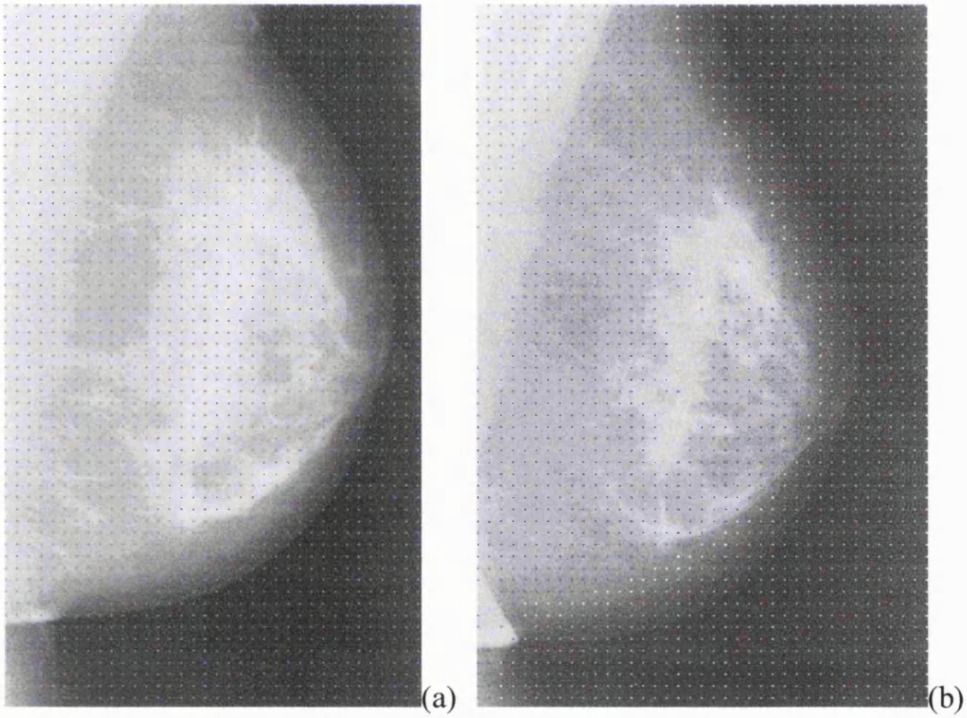


Figure 52: (a) The original mammogram, (b) textural segmentation of the pectoral muscle (in blue colour).



(a) (b)

Figure 53: (a), (b): The original mammogram pair.



(a) (b)

Figure 54: (a), (b): Texture clustering for the temporal pair in Figure 53.

6.5 Extension of registration to other modalities

6.5.1 X-ray, 3d-MRI data fusion

As discussed in chapter 2, data fusion between the main modalities for breast cancer imaging (MRI, US, Nuclear Medicine and X-ray) can assist the clinician in combining anatomical and functional information. This could be useful for diagnosis, surgical planning and therapy treatment. Recently, the possibility of "fusing" imaging data from digitised mammograms and contrast-enhanced MRI (Gd-DTPA), has been investigated showing very promising results [32, 33]. The objective is to correlate diagnostic information between the two imaging modalities (anatomical and functional). This could be done by registering CC and ML projections of the 3D MRI volume to the X-ray CC and ML mammograms and then relating mammographic features (e.g. microcalcifications) from the mammograms to the MRI volume. The correspondence of features between the two modalities is not obvious because the mammograms are compressed. In the rest of this section, we summarise the data fusion work [32, 33, 160].

The method consists of two important steps:

1. Projection of voxel contrast-enhancement. This is done in order to acquire an anatomically significant (and morphologically similar to mammogram image) projection of the 3D volume. This is done by:

    - Fitting a pharmacokinetic model of Gadolinium enhancement to each voxel in order to classify the rate and amplitude of enhancement. This is done to filter out the fat and most of the effects of moderately-enhancing parenchyma. This is illustrated in Figure 55, where the regions of high enhancement are favoured comparing to the low enhancement ones (e.g. fat).

    - Project each breast in the approximate CC and ML/MLO planes (Figure 56). However, these are projections of the uncompressed breast and in addition the MRI is acquired in the prone position. Figure 57, shows the "pseudo"

mammogram together with the original one. There is a clear correspondence both in the pathology (the cancer is clearly visible in both images) and in anatomical structures (e.g. vessels).

Subsequently the two "pseudo X-rays" in the CC and ML planes are registered to the CC and ML mammograms:

2. Non-rigid registration (matching) between the CE-MRI projections and the two mammogram views. This is achieved by extending the mammogram registration method presented in chapter 4 of the thesis to include the "pseudo X-rays". In summary:

   - The mammograms are "warped" to the shape of the projections

   - The matching process consists of two steps – first the breast boundary is used to "approximately" warp, and then internal landmarks complete the deformation.

Results of the registration technique are shown in Figure 58. The projected MRI volume (in the ML plane) is registered to the ML X-ray, using the boundary as well as internal landmarks. This way, the location of features that are mammographic features (e.g. microcalcifications) can be estimated in the 3d-MRI. This was exploited in [160, 161], in order to differentiate benign from malignant microcalcification clusters by observing the enhancement of the contrast agent in the MRI volume.

The main clinical applications include:

1) Women with dense breasts
2) Multi-focal disease and impalpable tumours
3) Detection of DCIS

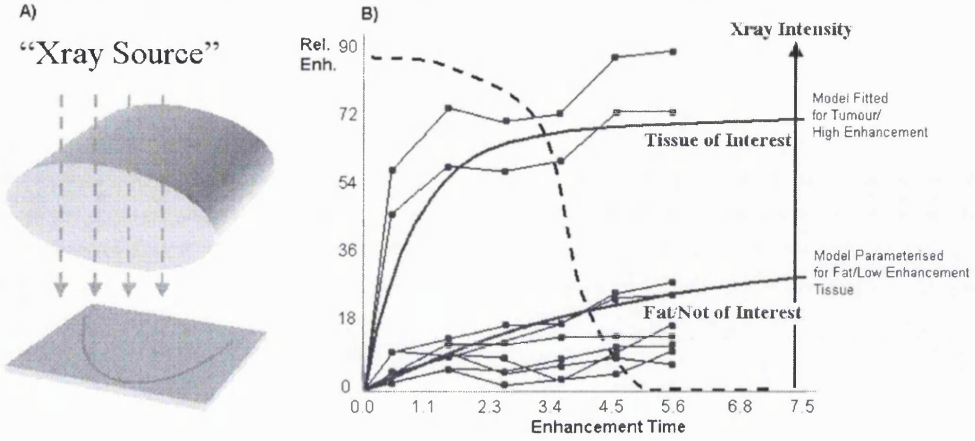


Figure 55: The pharmakokinetic model used for voxel classification.

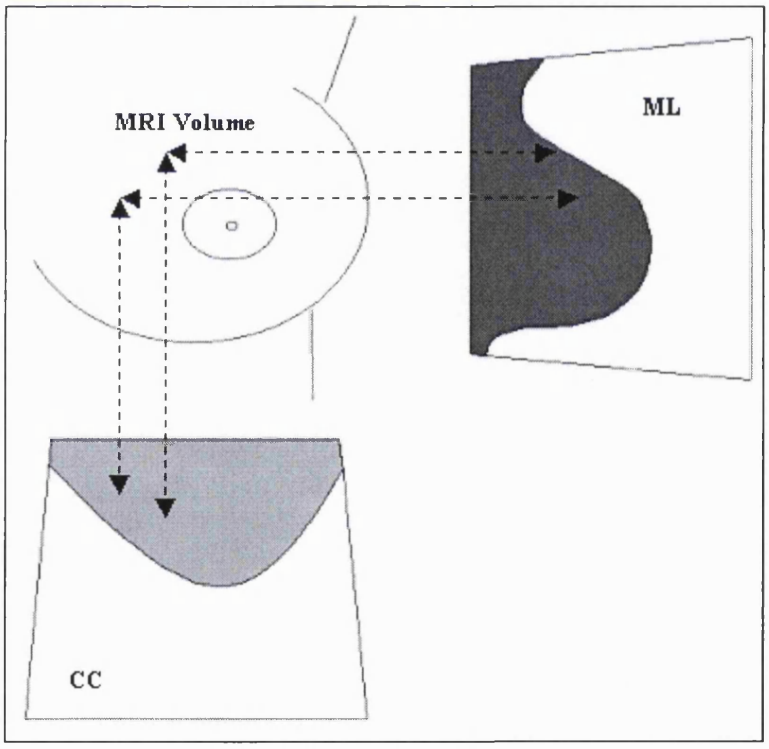


Figure 56: Data fusion between 3d MRI and mammography. In order to correlate the mammographic anatomical information with the MRI volume we need to establish correspondence between the 3d volume and both the cranio-caudal and the medio-lateral mammograms.

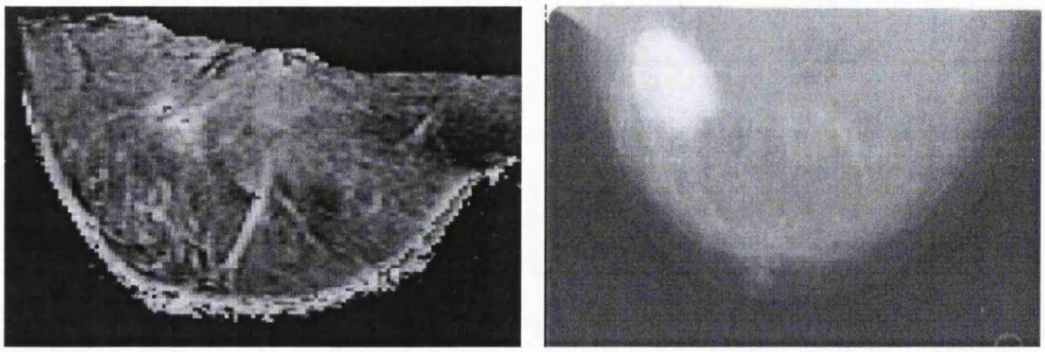


Figure 57: Visual comparison of the "pseudo mammogram" (on the left) and the mammogram show a good correspondence of features.

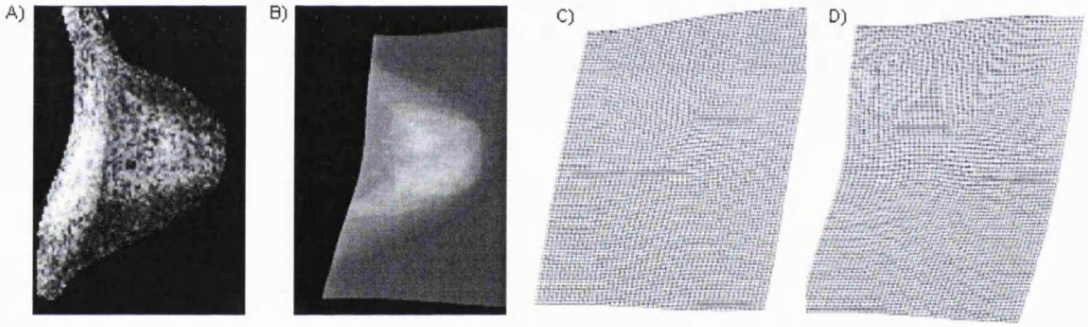


Figure 58: Registration of the ML projection of the MRI volume (A) with the ML mammogram (B), first using the boundary (deformation field shown in (C)) and then including internal landmarks (deformation field shown in (D)).

6.5.2 X-ray-Nuclear medicine data fusion.

Nuclear medicine modalities for breast cancer imaging, such as scintimammography (described in chapter 2), offer very good sensitivity for cancer detection, especially for younger women or for women that had previous surgery or implants in which case, mammography can be non-diagnostic. In Figure 59 (a) we can see the mammographic appearance of a young woman's breast. Due to the dense tissue the image appears bright and as a consequence the contrast between dense structures (e.g. density and cancer) is very low. Figure 59 (b), is a typical scintimammogram of a young woman that has developed cancer (there are two masses that are highlighted in the scan).

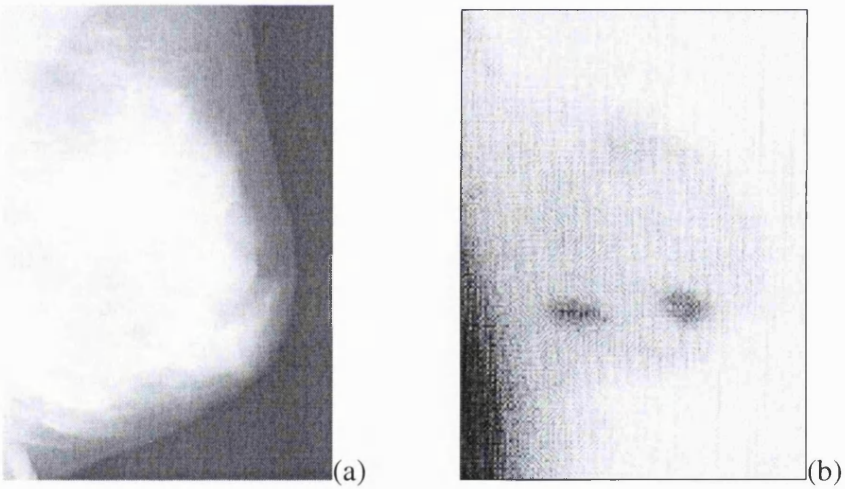


Figure 59: Mammographic appearance of a young woman's breast, (a): A mammogram where the dense tissue makes diagnosis extremely difficult and (b): a nuclear medicine scan where only the pathological regions appear.

The benefit of registering X-ray mammograms to scintimammograms are obvious: Firstly it would enable the clinician to combine the rich anatomical representation of the mammogram with the detected pathology (in the scintimammogram). In addition, using the registration transformation one can estimate the location of the pathology in the

"dense" mammogram in order to assess retrospectively the possibility for a mammogram-based diagnosis.

A scintimammogram is aquired with the woman lying in the prone position. As shown in Figure 59, the low contrast in the breast edge could pose a problem in detecting the breast outline from the scintimammogram (for registration). However, nipple markers can be used in scintimammography and this can facilitate boundary-based registration with mammograms.

6.6 Future work on HRT sequences

In chapter 5 of the thesis, we demonstrated a framework for quantitative image analysis of HRT sequences. To benefit from this analysis a large scale experiments is necessary in order to study the density history of a large number of women that developed cancer while receiving HRT. Effectively, we'll register temporal sequences for each woman and calculate density changes over time (before and after the cancer appeared). In addition, similar studies should be made for women that did not develop any pathology while on therapy. For both categories the following measures can be calculated:

- Global density changes over time

- Local changes in density of specific areas

- Density changes per breast quadrant. This would require the employment of a 3D uncompression model [162] where the location of density change could be estimated in the 3d representation of the breast from the CC and ML mammograms.

An exciting challenge would be to build a model of "normal" density variation against a "pathological" one based on a large number of data. A first step could be to build a geographical distribution model of "abnormal" density changes, based on a large number of HRT mammogram sequences.

6.7 Final statement

This thesis has described work to register temporal mammograms for a more efficient comparison of changes. A two-stage algorithm has been proposed to improve the registration based *only* on the boundary. Moreover, the issues related to HRT use have been discussed and an image analysis framework for assessing both *global* and *local* changes in tissue density between temporal HRT mammogram pairs has been suggested. Results of the methods proposed have been included in the main text of the thesis as well as in a separate Gallery of results (Appendix E).

Appendix A: Breast anatomy, physiology and pathology

1 Anatomy and Physiology of the Female Breast

The female breast is a well-differentiated apocrine gland [163] that originates in the ectoderm and secretes milk during lactation. Each breast is located between the sternum and the mid-axillary line and anterior to the pectoral muscles. Figure 1 illustrates the basic structures of the breast. From the top to the bottom we can see the glandular tissue, the Cooper's ligaments and the adipose (fatty) tissue (represented in yellow colour) together with the vessel network.

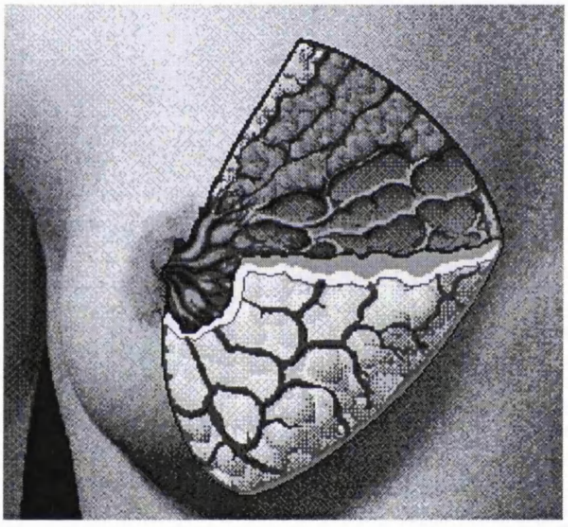


Figure A1: Basic Anatomical Features of the Mature Female Breast

The most important part of the breast is the glandular tissue, which is often referred to as fibroglandular tissue because of its indistinguishable coexistence with fibro-adipose tissue. The glandular tissue is very dense and consists of 15-20 lobes. Each lobe is drained by one lactiferous duct and all ducts converge to the area around the nipple, called areola. Each lobe is composed of a network of ducts and lobules that brunches in the direction towards the sternum. The end units of these networks are called terminal duct-lobular units. Most benign changes and almost all breast cancers arise within the terminal duct-lobular unit [22]. The milk is produced in the end lobules and is drained radially through the ducts to the nipple [27].

The breast tissue is attached to the overlying skin by straps of fibre called Cooper's ligaments. They are scalloped strings of fibrous connective tissue, located throughout the breast to support the glandular tissue together with the duct the blood vessels, the lymph channels and adipose tissue. The latter is found in the outer portion of the breast, surrounding the glands and filling the intervals between the lobes of the breasts. Fat tissue is a metabolically active tissue that acts as a reserve supply of energy by storing fat and releasing it in response to a variety of nervous and hormonal stimuli [164]. It also acts as an insulator to help maintain body temperature and acts as a protective padding in certain areas. Fatty tissue also helps smooth out the contours of the body.

Finally, a network of veins and arteries carry blood to and from the breasts while lymph fluids wash away waste products, dead cells, and other debris. These fluids flow along lymph vessels, which empty into lymph nodes. There are approximately 35 lymph nodes around the breast, most of which are located in or near the armpit [164]. The axillary lymph nodes receive approximately three quarters of the total lymph drainage. This is the cause of very frequent tumour metastases to these nodes as the lymph, along with other debris, can carry cancerous cells [8].

2 Pathology

There are two major categories of aberrations of normal breast development, benign changes and malignant pathology. The most common benign disorders are:

Fibroadenomas: They are described as fibrous overgrowth, developed from a whole lobule and not from a single cell. They account for 13% of all palpable symptomatic breast masses, and for 60% in women aged 20 or less. In clinical practice they are diagnosed by clinical examination in combination with ultrasonography and fine needle aspiration cytology (in older patients they have characteristic mammographic features, when they calcify). The decision upon excision or not depends on the size of the fibroadenoma and the age and history of the patient.

Cystic disease: Cysts are distended and involuted lobules and appear as smooth and discrete breast lumps and account for 15% of all discrete breast masses [27]. They exhibit a characteristic halos in mammography but are usually diagnosed by ultrasonography because they don't produce any echoes. The diagnosis is established by fine needle aspiration.

Sclerosis: During stromal involution there are many possible disorders including the development of localised areas of excessive fibrosis or sclerosis. The mammographic appearance of sclerosing lesions can mimic that of cancer, leading in diagnostic problems.

Epithelial hyperplasia: This is an increase in the number of cells lining the terminal duct lobular unit. If, in addition these cells are atypical the condition is called atypical hyperplasia, which unlike any other benign disease is associated with an increased risk of subsequent breast cancer.

Finally, breast cancer is a malignant disease which originates from the epithelial duct cells. The term cancer describes a group of diseases in which symptoms are due to unrestrained growth of anaplastic cells (cells that grow without form or structure) in one of the body organs or tissues. Research in this field has demonstrated the correlation between cancer and tobacco, natural constituents of food, sexual and reproductive history, occupational hazards, alcohol and food additives.

If the cancer cells remain within the terminal duct lobular unit, they are classified as in situ or non-invasive. Ductal carcinoma in-situ is the most common type of cancer in this category, accounting for 17% of screen detected cancers [22]. This type of disease is characterised by the appearance of microcalcifications. The other type of in-situ cancer is less common (1% of screen detected cancers) and is known as lobular carcinoma in-situ. It is very important that the cancer is detected at an early (non-invasive) stage, thus reducing the possibility of spreading in the surrounding tissue through surgery.

In many cases cancer infiltrates the surrounding tissue, leading to the condition known as invasive carcinoma. Depending on the origin of the cancer, they are characterised as ductal or lobular. At this stage surgery, chemotherapy and hormonal treatment are the most effective weapons against the spread of cancer. If the cancer infiltrates surrounding lymph nodes it is spread to other parts of the body like the spine, the liver and the brain and becomes very difficult to control.

The main problem with imaging breast cancer is that affected lesions are very small and usually hidden among the glandular tissue. For that reason newer modalities (e.g. Breast-MRI or Nuclear Medicine) are being used in conjunction with mammography, in an effort to detect early malignant changes such as increased vascularity.

3 Risk factors for breast cancer

Some of the most important risk factors for breast cancer include:

- Age is a risk factor, since women over 50 years old are more likely to have breast cancer than women less than 50 years.
- Demographic characteristics, family history and genetic predisposition.
- Previous breast cancer or benign breast disease.
- Mammographic pattern (e.g. P2 DY high-risk patterns [8]).
- Oral contraceptives, Hormone replacement therapy and irradiation

4 Breast Structure and Parameters Important for Imaging

The breast begins to develop during and after puberty. The major changes occur during pregnancy preparing the breast for milk secretion. In the fully developed breast the glandular tissue has developed to its maximum size and at that stage it is very difficult to image the breast with x-ray mammography. This is because the dense fibroglandular tissue attenuate a large amount of x-rays resulting in a low contrast, high intensity region

in the display. Finally after menopause the glandular tissue is gradually replaced by adipose tissue. This procedure is called involution and is characterised by the variability in its rate between different women and regions of each breast [27, 163].

For all the reasons mentioned, it is clearly very important to image the glandular tissue of the breast. Nevertheless glandular tissue is a dynamic structure. Hormonal variations related to menstrual cycles, pregnancy and lactation, in combination with factors such as stress, changes in nutrition or bodyweight, lead to everyday changes in the size and subsequently the mammographic appearance of the glandular tissue.

Breast Asymmetry is a very important parameter and it is often useful to prove such a condition by imaging the breast. Asymmetry between breasts is not considered as a highly indicative sign of malignancy [8], but it is still a very important mammographic sign to detect.

Architectural distortion is an important feature for especially for x-ray mammography, and is characterised by a shift in an area of breast tissue. Thus a part of the breast tissue appears to be displaced. Architectural distortion is considered as a sign of malignancy.
Calcifications are strongly related with the presence of breast cancer. Macro-calcifications are structures greater than 1mm in diameter, and are indicative of benign breast disease. On the other hand micro-calcification is associated with malignant disease, but in some case it can also appear in benign changes. Many researchers from the field of computing and image processing have worked in the automatic detection of calcifications in digitised mammograms [165, 166].

Appendix B: Wavelet segmentation of internal regions in mammograms

1. Introduction

In this Appendix we present the multi-scale segmentation method used in this thesis for detecting internal structures in mammogram pairs. This method was developed by C. Behrenbruch and all the material in this Appendix is from [160].

2 Overview of Wavelet Concepts

Various spatial filtering techniques based on localised Fourier operators and trigonometric transforms have been used for image feature detection [167-170]. The inherent limitation of these approaches lies in the need to define a large number of Fourier windows (in a multi-resolution context) in order to produce a maximally sensitive feature detector. In order for these techniques to work robustly, there is a need to be able to adapt the spatial localisation properties of filters as meaningful descriptors with respect to image content. Most windowed Fourier/Gabor techniques can do this only to a limited extent.

As an alternative, we extend the notion of a multi-resolution analysis to wavelets via the use of dilation and scale operators in the frequency domain:

$$s(x) \alpha \; s\left(\frac{x}{a}\right)/a \Leftrightarrow \bar{s}(\omega) \alpha \; \bar{s}(a\omega) \qquad (1)$$

$$s(x) \alpha \; s(x-b) \Leftrightarrow \bar{s}(\omega) \alpha \; \bar{s}(\omega)e^{-i\omega b} \qquad (2)$$

where the frequency domain is re-parameterised in terms of scale and dilation parameters, a and b. This results in a class of integral transforms – "wavelet transforms" – that are represented as a convolution of a given function or sequence $f(x)$ with a kernel function (filter):

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\overline{\psi}\left(\frac{x-b}{a}\right) \qquad (3)$$

and the transform representation over all scales as:

$$Wf(a,b) = \int_{-\infty}^{\infty} f(x)\overline{\psi}_{a,b}(x).dx$$
$$= \int_{-\infty}^{\infty} f(x)\frac{1}{\sqrt{a}}\overline{\psi}\left(\frac{x-b}{a}\right)dx \qquad (4)$$

This transform is known as the "continuous wavelet transform" (CWT) which completely describes the decomposition of a function via "wavelet" functions, much the same way as a Fourier decomposition can describe an arbitrary (continuous) function by sinusoidal elements.

Referring back to equation (1), in the case that the transform of s is localised around $\omega_0 = 0$, the dilated function $s(a\omega)$ is still localised around $\omega_0 = 0$. The dilation operator does not change the translation in the frequency domain, only the bandwidth of the filter. Therefore it is a requirement that s does not contain the zero frequency or is of zero mean:

$$\bar{s}(0) = 0 \Leftrightarrow \int_{-\infty}^{\infty} s(x).dx = 0 \qquad (5)$$

This definition is presented here as it explains better the use of the term "wavelet" (as a filtering concept) since the dilation operator must contain some oscillation and cannot be a mere "bump". Therefore Gaussian-type convolution [171, 172] is not a wavelet filter, however (for example), there is a relationship with Laplacian scale-space [173, 174], difference of Gaussian analysis [175], and the time-varying solution of the heat equation [176]. An example of one of the simplest wavelets – the "Morlet" – is shown in Figure B1 as the modulation of a cosine with a Gaussian function [177].

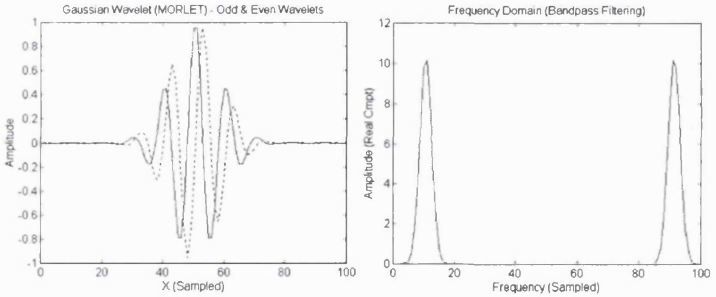


Figure B1: A simple wavelet function (Morlet) generated by modulating a cosine with a window function (Gaussian). The corresponding frequency domain representation is illustrated to show the resulting filter transfer function.

Equation (5) also defines the invert ability condition of the wavelet, which leads to the inverse wavelet transform for reconstruction:

$$f(x) = \frac{2}{C_\psi}\int_0^\infty \frac{da}{a^2}\int_{-\infty}^{\infty} W_\psi f(a,b)\left[\frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right)\right].db \qquad (6)$$

The previous equation invites the observation that the CWT is an information inefficient representation as a 1D function is parameterised by the 2D $f(a,b)$. This redundancy can be compensated by subsampling the spatial-frequency representation of the wavelet basis $W_\psi f(a,b)$ at intervals $\{ a = a_m \mid m \in Z \}$, $\{ b = b_n \mid n \in Z \}$. This leads, in turn, to the work of Mallat [178, 179], which defines the "discrete wavelet transform" (DWT), which can be effectively implemented as a filter bank. The result is a dyadic (factor of two) signal decomposition where the dilation is performed via successive convolution and decimation with quadrature filters (QFs) H and G for a discrete sequence $S^j(n)$ resulting in an *approximation* A^{j-1} and a "detail" component, D^{j-1}. Similarly, reconstruction is performed via an adjoint convolution and upsampling. This is illustrated in Figure B2.

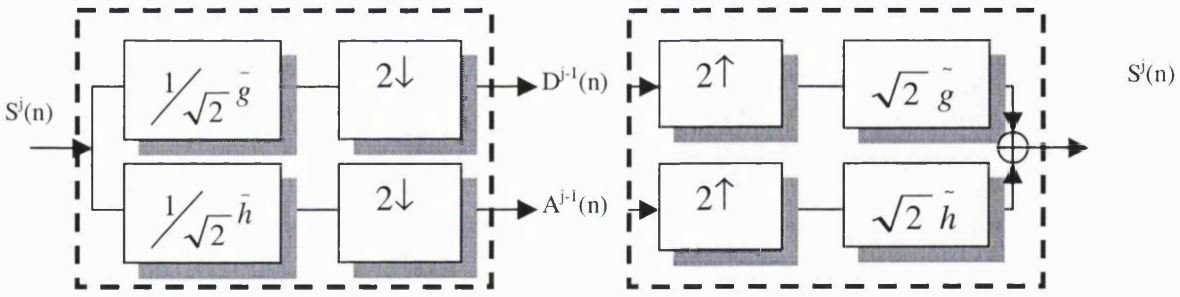


Figure B2: The process of dyadic wavelet decomposition involves convolution and decimation of a sequence with a quadrature filter pair H and G (represented as h,g – the coefficient sequence) into an approximation A and a detail D component. Reconstruction of the sequence is via convolution with the adjoint pair and resampling.

The extension of Figure 2 to several scales (successive convolution/decimation) results in the concept of "multi-resolution analysis" (MRA). Recursive filtering can be used to decompose a function or sequence into approximations at different scales, by effectively

changing the successive support length of the filter kernels. An example of such an approximation scheme is shown in Figure B3 where a 1D cross-section of a ROI in an image is successively decomposed (a_1, a_2, a_3...).

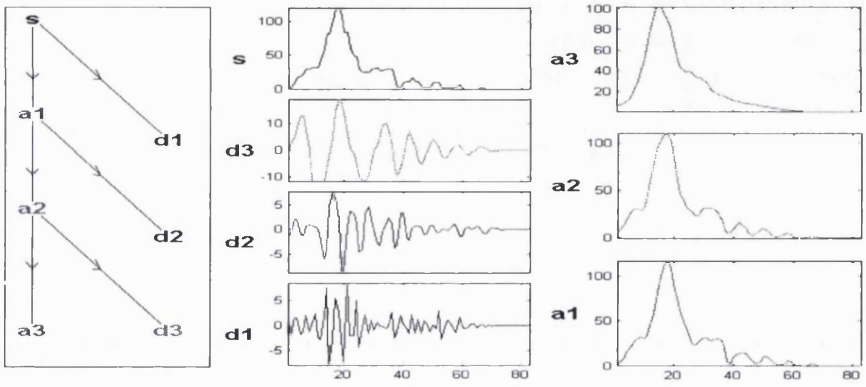


Figure B3: An example discrete wavelet analysis showing the successive approximation of the MRA into detail and approximation components (using the well-known Daubechies wavelets [180], filter length 4).

Similarly, for the MRA case, adjoint convolution and upsampling can reconstruct the original function or sequence from a decomposition across many scales. If various projections are not used in the reconstruction process, it is possible to regenerate a smoothed approximation or, in other words, "denoise" a signal [181, 182]. If an information cost function is used to select *only* those projections that represent components in the function or sequence at a specific scale then this framework can be used for feature detection. This idea is developed further in section 7. It is worth noting that alternative (efficient) wavelet scale-space concepts continue to be developed and may be equally applicable to feature detection.

3 Wavelet Packets

The discrete wavelet transform as a dyadic filtering implementation provides a multiscale foundation for image analysis. However looking at the wavelet decomposition in terms of filter banks, there is an obvious question to be asked. What if the projection space

generated by the filter bank "tree" was extended to the equivalent high-pass case, as illustrated in Figure B4.

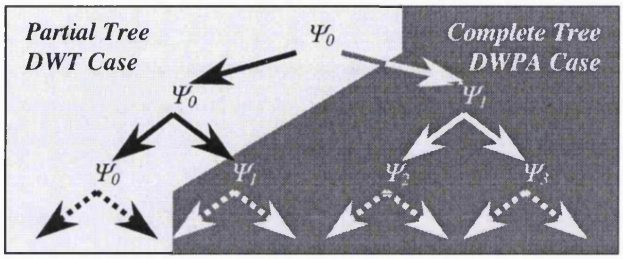


Figure B4: A tree representation showing the additional decomposition components (quadrature filtering) of wavelet packets, compared with the basic DWT case.

It turns out that this reconfiguration of the basis subspaces leads to "discrete wavelet packet analysis" (DWPA) [110, 183] which essentially produces particular linear combinations (superposition) of wavelets, resulting in a richer library of wavelet projections (called "packets") with which to analyse a function. Additionally, DWPA forms subspaces that retain the orthogonality, smoothness and localisation properties of the dyadic DWT, but with better spatial localisation characteristics.

The improved spatial characteristics are demonstrated in an example in Figure B5, where an (approximate) impulse function is decomposed using both wavelets and wavelet packets. When viewed as a tiling of information cells across the space-frequency plane, the DWT analysis correctly localises the peak of the function in the high frequency elements, but is forced to include poorly localised, low frequency components as well. Although the analysis equivalent for the wavelet packet case does not contain all tilings, it demonstrates that it is possible to reconstruct the signal using almost a single wavelet packet. For this reason, we have used DWPA as the foundation of our feature detection.

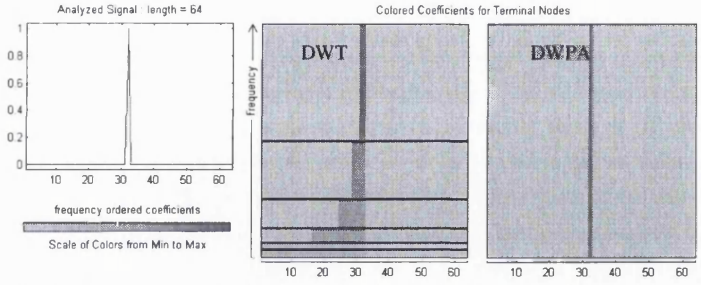


Figure B5: A comparison between DWT and DWPA representation of an (approximate) impulse function. The DWT case (left) correctly determines the location of the peak for the high-frequency components but also includes low frequency components with poor spatial localisation. The DWPA case (right) demonstrates much better spatial localisation and produces an optimal representation of the impulse.

The starting point for defining wavelet packets mathematically lies in the interpretation of the tree diagram of Figure B4. Using H and G to represent the equivalent high and low pass filter components of a set of conjugate quadrature filters, the following recursive set of functions can be defined:

$$\psi_0 \overset{def}{=} H\psi_0; \int_R \psi_0(t).dt = 1 \tag{7}$$

$$\psi_{2n} \overset{def}{=} H\psi_n; \psi_{2n}(t) = \sqrt{2}\sum_{j\in Z} h(j)\psi_n(2t-j) \tag{8}$$

$$\psi_{2n+1} \overset{def}{=} G\psi_n; \psi_{2n+1}(t) = \sqrt{2}\sum_{j\in Z} g(j)\psi_n(2t-j) \tag{9}$$

The function ψ_o (the "root") is uniquely defined and is identified with wavelet generation function ϕ in [178]. The function ψ_1 is the "mother wavelet" associated by convolution with H and G. The descendent functions (tree "leaves") ψ_n are analogous to iterated forms of the mother wavelet as in the DWT, however in this case, the evolution of ψ_n is over the complete tree. The collection of these functions for $n = 0,1,\ldots,m$ makes up the wavelet packets associated with a particular choice of quadrature filters. It is also apparent that in this way, wavelet packet analysis is a generalisation that includes the DWT (i.e. only part of the tree is used), as shown in Figure B4.

Clearly, the easiest way to use wavelet packets is in a fixed scale context. In this way, they are somewhat analogous to analysing a function or sequence with a DWT at a fixed scale or level, albeit with a different set of filter projections. In our case, we use wavelet packets in a multiscale context, which effectively means that the approximation can be performed using projections anywhere in the decomposed basis. For further details of the concepts of wavelet analysis, we refer to [184] and for a very useful overview of signal processing with wavelet packets (as well as a good treatment of the software implementation issues) [185].

4 Best Basis Selection

The multiscale wavelet packet representation provides a rich set of filtered projections within a wavelet basis at different scales. Using wavelet packets for feature detection involves ordering these subspaces by information content to produce a scale-space "stack" for analysis. This approach is called "best basis analysis" [186]. Before constructing such a scale space, it is necessary to define the notion of an information cost function that is used to rank each representation in terms of the contribution to approximating a sequence. To do this, we use a real-valued functional M, satisfying the condition:

$$M(u) = \sum_{k \in Z} \mu\left(\left|u(k)\right|\right) \; ; \mu(0) = 0 \qquad (10)$$

Here μ is a real-valued function defined on $[0,\infty)$. If we assume that $\Sigma_k \mu(|u(k)|)$ converges absolutely, then M will be invariant to rearrangement of the representation of the sequence u. We can then define a functional M_x on the set of bases B for each $x \in X$:

$$M_x : B \to R; B \to M(B^* x) \qquad (11)$$

which is referred to as the *M-information cost* of a sequence x in the basis B. Another way of looking at this, is to evaluate a transformation of $f(x)$ to $g(x)$ via an operator (in this case convolution-decimation) as a distance measure with respect to a transformation parameter λ.

$$\left\| f - g \right\|^2_{L_2(I)} + \lambda \left\| g \right\| \qquad (12)$$

evaluated as:

$$\|f-g\|^2_{L_2(I)} \stackrel{\Delta}{=} \left(\int_I |f(p)-g(p)|^2 dx \right)^{1/2} \tag{13}$$

This enables each approximated component of the wavelet packet decomposition to be compared and subsequently ranked by information content. There are many different kinds of information cost functionals that can be used. For example, a simple threshold ε based on coefficients in a wavelet subspace could be used to order the contribution of each packet to the complete reconstruction of a sequence, as shown in equation (15).

$$\mu(\omega) = \begin{cases} |\omega|, if\, |\omega| \geq \varepsilon \\ 0, if\, |\omega| < \varepsilon \end{cases} \tag{14}$$

Another useful approach is the entropy of a representation, such that a cost function $u=\{u(k)\}$ is evaluated by:

$$H(u) = \sum_k p(k) \log \frac{1}{p(k)} \tag{15}$$

where $p(k)=|u(k)|^2/\|u\|^2$ is the normalised energy of the k^{th} element of the sequence. This functional can be evaluated in a minimisation context and used to rank a set of decompositions.

One of the difficulties with using conventional entropy measurements is defining a probability density function p in a way that is appropriate to each representation. Practically, this means selecting a histogram sampling density that meaningfully captures the statistical characteristics of the sequence. Clearly, in a multi-scale context where this distribution may vary greatly (and with different sampling resolutions by virtue of decimation), this is not a trivial task. To circumvent this problem, we have chosen to use an approximation to entropy – *Approximate Entropy* (ApEn), which is a useful measure of signal regularity [45,46]. The idea is that wavelet packet decompositions that smoothly approximate the original sequence will have a higher degree of regularity, whilst

equivalent high-pass filtered components that reflect the signal noise and will be more fluctuating in nature, exhibiting greater irregularity. ApEn is defined as follows:

$$\Phi^m(r) = \frac{1}{N-m+1}\sum_{i=1}^{N-m+1}\log C_i^m(r) \qquad (16)$$

where given a positive integer N and nonnegative integer m, with $m \leq N$, a positive real number r, and a sequence of real numbers $u:=(u(1),u(2),\ldots,u(N))$. The distance between two sequences $x(i)$ *and* $x(j)$, where $x(i)=(u(i),u(i+1),\ldots,u(i+m-1))$ is defined by $d(x(i), x(j)) = \max_{p=1,2,\ldots,m}(|u(i+p-1)-u(j+p-1)|)$. $C_i^m(r)$ is the number of $j \leq N-m+1$ such that $d(x(i),x(j)) \leq r$, normalised with respect to $N-m+1$. Thus, $ApEn(m,r,N)(u)$ may be interpreted as a measure of the frequency at which number sequences of length m occur compared with sequences of length $m+1$. High values of the ApEn measure imply a highly random nature, whereas low values suggest a relatively smooth sequence.

Therefore, if we apply the ApEn measure to each wavelet packet decomposition of a sequence, we can order each projection by the contribution it makes to completely reconstructing the sequence. Starting from the lowest approximate entropy packet decompositions and progressively adding detail (higher ApEn), one can construct a scale-space using multi-scale wavelet packets, despite the fact that different scales (depending on whether they are equivalent low-pass or high-pass filtering combinations) may be ranked counter-intuitively by virtue of the level of decomposition. An example of such a scale-space is shown in Figure B6, where a small portion of an X-ray image is decomposed into a wavelet packet scale-space stack (2D sequences are discussed more thoroughly in the next section).

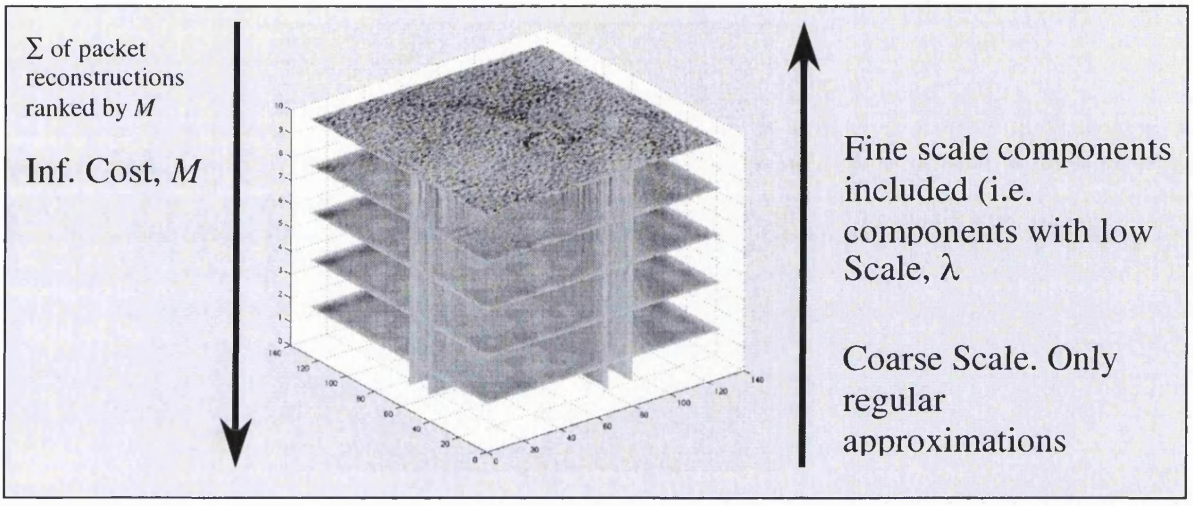


Figure B6: An example scale-space illustrating cumulative addition of image detail (via packet reconstructions) using an information cost function. The bottom of the scale-space contains a quite smoothed approximation to the image, whilst the top of the stack is a perfectly reconstructed image (i.e. all packets have been used).

5 Choice of Wavelet Packet

Because of the superposition of wavelets, wavelet packets provide a large range of multi-scale representation with which to create a scale space. For our work, we have chosen a class of existing wavelet bases, called "Coiflets", which have a number of useful properties that make them ideal for detecting regions of interest in medical images. In particular, we wanted a wavelet basis that would be sensitive to sharp, spatially localised features such as tumour spiculations, microcalcifications (in X-ray images) and small regions of focal enhancement in MRI (such as for ductal carcinoma). Coiflets, have the following useful properties:

- They have a high number of vanishing moments for a given support length (i.e. good frequency localisation as well as spatial localisation)
- They are almost symmetrical
- They have compact support
- They are orthogonal (indeed a bi-orthogonal form is also available)

Low-order Coiflet wavelet packets are useful for approximating "bumps" without over-smoothing, hence they have good edge-preservation characteristics. Higher-order packets have broad frequency characteristics making them useful for noise approximation. These properties make Coiflets very good for generating the kind of scale-space shown in Figure 6, as a good spread of detail is ascribed to both the smoothly approximated region of the scale-space and the high-frequency (noise) components. In the case of features that have a broad Fourier localisation (for example, sharp changes in intensity due to edges or highly localised features), the Coiflet scale-space provides good discrimination from smoother components. Some example Coiflet wavelet packets are shown in Figure B7 (generation and scaling functions shown, as well as example higher-order packets). For more information on the generation and properties of Coiflets we refer to [180]. Additionally we are considering the development of customised biorthogonal wavelets based on the "lifting scheme" [189], which could be tuned to the morphology of specific features or feature characteristics.

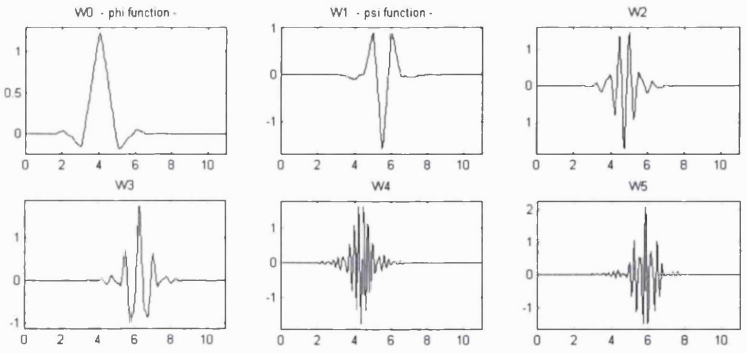


Figure B7: Example Coiflet wavelet packets (C2) including the generator/scaling functions (W0/1) and higher order members.

6 Extension of Wavelet Packets to Images (2D)

Our treatment of wavelets, wavelet packets and scale-space construction via information cost functions (best basis analysis) has so far been limited to the basic 1D case. However, for analysing images, these concepts need to be extended to higher dimensions.

Fortunately, most conjugate quadrature filtering algorithms extend quite neatly to higher dimensions via the tensor product of 1D basis elements:

$$b(x) = b(x_1,...,x_d) = b_1(x_1)\cdots b_d(x_d) \qquad (17)$$

This tensor product basis approach is termed "separable" because they can be factored across integrals to obtain a sequence of d one-dimensional problems by treating each dimension individually. The equivalent low- and high-pass filters H and G in the definition of the wavelet packet decomposition are quadrature filters with corresponding dual (adjoint) filters that can be used for reconstruction. Therefore the tensor product approach to extend the wavelet packet analysis to higher dimensions naturally involves the tensor product of 1D sequences (filter coefficients), h and g. We define the tensor product of quadrature filters on the bivariate sequence $u=u(x,y)$ as:

$$\begin{aligned}(H \otimes G)\, u(x,y) &= \sum_{i,j} h(i)g(j)u(2x-i,2y-j) \\ &= \sum_{i,j} h(2x-i)g(2y-j)u(i,j)\end{aligned} \qquad (18)$$

Similarly, for the adjoint (reconstruction) case, the tensor product $(H \otimes G)^* \equiv H^* \otimes G^*$ is defined by:

$$(H^* \otimes G^*)u(x,y) = \sum_i \sum_j \bar{h}(2x-i)\,\bar{g}(2y-j)u(i,j) \qquad (19)$$

These definitions generalise to an arbitrary dimension d for a filter operator $F^i \in \{H,G\}$ for $i=1,\ldots d$:

$$\begin{aligned}(F^1 \otimes \cdots \otimes F^d)u(x) &\overset{def}{=} \sum_{k_1} \cdots \sum_{k_d} f_1(k_1)\ldots f_d(k_d)u(2x_1-k_1,\ldots,2x_d-k_d \\ &= F^1_{x_1} F^2_{x_2} \ldots F^d_{x_d} u(x)\end{aligned} \qquad (20)$$

A visualisation of equation (20) for $d=2$ is shown below (Figure B8).

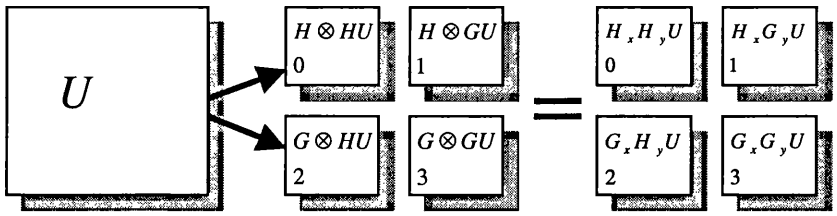


Figure B8: The implementation of wavelet packet analysis for $d=2$ is performed by directional convolution and decimation for each dimension of the image.

This approach is analogous to subb and coding, a relatively conventional approach to image decomposition. In terms of implementation, this means that the image is convolved (and decimated) with the quadrature filters in each dimension in a "quadtree" fashion. As an aside note, the actual implementation of tensor wavelets requires some care with respect to the issue of mixing scales via tensor products [184].

7 Feature Extraction

Now that we have defined scale-space in terms of "best-basis" ordered reconstructions from wavelet packets, it is possible to perform feature extraction. The extraction process involves region growing the scale-space (as shown in Figure B6) and identifying regions that persist through many levels of scale.

Unlike the "extremum" stack concept of Koenderink [171] or the conventional CWT scaleogram [190], our scale-space construction does not guarantee continuity through the entire stack. This is the result of an information cost function to order the wavelet packet construction. Therefore, we are interested in regions that may be well localised across several scales, but not necessarily in a congruent fashion across *all* scales. The most significant n regions are treated as "features" and are labelled by the highest level of scale reached by region growing/merging. A visualisation of this effect is illustrated in Figure B9 where a small portion of an X-ray image containing microcalcifications (sharp, high-intensity features) is visible as distinct regions propagating through the stack.

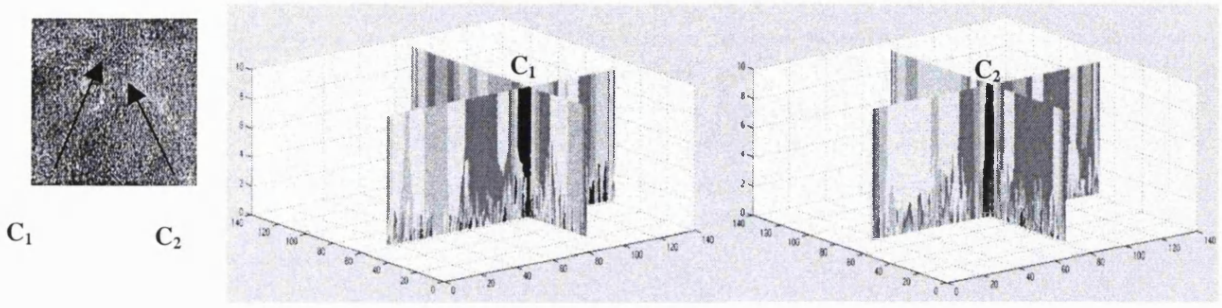


Figure B9: A small section (140x140 pixels) of a 30μm digitised mammogram containing two calcifications is decomposed into the wavelet scale-space (as shown in Figure 8). The calcifications are detectable as a well-localised propagation through the sliced scale-space (C_1 and C_2). Note: The scale-space has been inverted to better show the shape preservation from coarse to fine scales.

The region growing starts with a collection of seed pixels at the low-scale part of the stack. A merging operator is applied to adjacent pixels l_{i+1}, l_{i-1}, l_{j+1}, l_{j-1} and l_{k-1},l_{k+1} in the scale-stack volume $V(x,y,\lambda)$ using the following criterion:

$$\Delta E = \frac{L_i L_j L_k}{L_i + L_j + L_k} \| \frac{\sum_{n \in X} V(n)}{L_i} - \frac{\sum_{n \in Y} V(n)}{L_j} - \frac{\sum_{n \in \Lambda} V(n)}{L_k} \|^2 - \Gamma \qquad (21)$$

and

$$\Delta E \begin{cases} < 0, l_i, l_j, l_k & merge \\ >= 0, l_i, l_j, l_k & unchanged \end{cases} \qquad (22)$$

where

$$L_i = \sum_{n \in l_i} L(n), \quad L_j = \sum_{n \in l_j} L(n), \quad L_k = \sum_{n \in l_k} L(n) \qquad (23)$$

and $L(n)$=1 and L_i gives the vector of intensities of l_i. This process is repeated for a fixed threshold value Γ until no more merges are possible (i.e. the region is completely grown). Γ is then increased until either the desired number of regions is produced (the m most significant regions) or the number of regions converges. This results in a segmentation of the scale-space that can be visualised as an overlay on the image under analysis (shown in Figure B10).

Figure B10: Segmentation results for the small region of a mammogram illustrated in the previous figure. There are several features in this image (left) including a small curvilinear structure and two clusters of calcifications (circled). The segmentation (right) has produced several regions of interest whilst being relatively impervious to the noisy characteristics of the image.

Once the regions have been segmented, they are fitted with a cubic spline contour and the image labelled into generalised areas of "region of interest" and "noise". The enables a histogram of pixel values to be assigned to "features" or regions of interest – ROIs - (up to a certain scale) and "noise", which has a different statistical distribution of intensities. With these estimations for feature and noise, we utilise the region competition algorithm of Zhu [191] to slightly refine the contour boundaries of the detected features, as a small amount of smoothing has transpired during the approximation process. This refinement step is not strictly necessary, although it produces feature contours that have a higher degree of complexity and better reflect the actual morphology (for example the fine spiculations of a tumour). An example of a completely segmented mammogram image is shown in Figure B11. It is this collection of features that is used as the basis of the internal landmark matching.

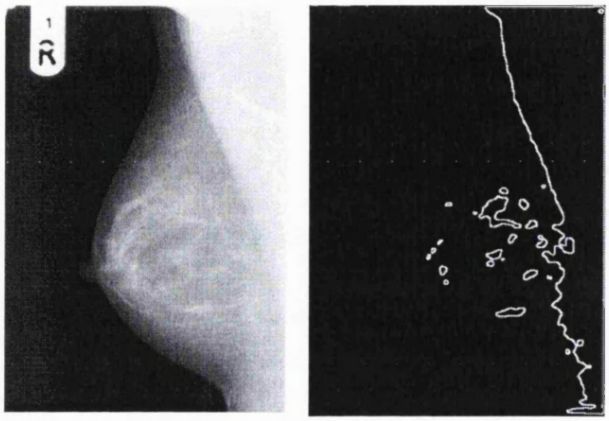


Figure B11: An example of wavelet-based feature detection and segmentation of a mammogram of a patient with very dense breasts. This is a good example as it shows image scale-space can be used to produce a limited number of high quality landmarks representing visually salient features.

In order to define landmarks for the RBF deformation, we ignore the pectoral muscle region and the radiological tag and since our features are based on closed contours grown from the scale-space stack, the centroids of the features are defined as candidate RBF landmarks. Clearly, this introduces a small amount of error into the registration, however the approximation framework largely compensates for this in the completion registration step and is also influenced by the number of landmarks used.

Appendix C: The h_{int} representation

1 Introduction

The h_{int} representation was discussed in several parts of the thesis. In this Appendix we summarise the basic concepts of the h_{int} representation for normalising mammogram *focusing on the method for the calculation of* h_{int}. A complete account can be found in [4]. In the h_{int} image, each pixel represents the thickness of 'interesting' (non-fat) tissue between the pixel and the X-ray source and the image surface provides anatomical information about the breast. In Figure C1, we can see an h_{int} surface; A cyst is visible at the bottom left and appears as a significant hill comparing to the surrounding tissue. Using the h_{int} representation, Highnam and Brady modelled and removed the effect of the scattering and extra-focal radiation thus creating a general framework for mammogram normalisation.

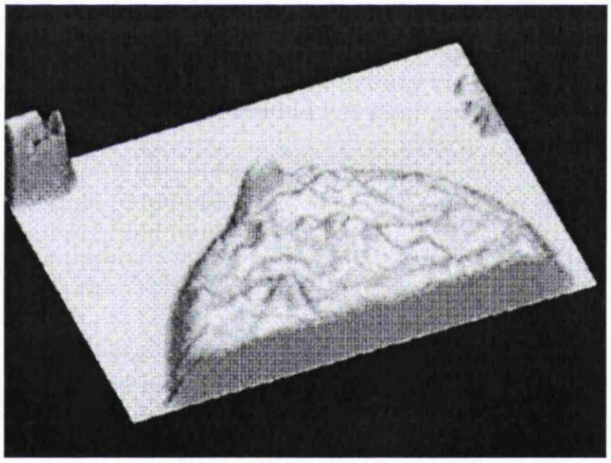


Figure C1: The h_{int} representation of a mammogram where a cyst is visible at the bottom left as a significant hill in the anatomical landscape.

As far as temporal mammograms are concerned, using the h_{int} representation, we can compensate for the non-rigid relationship of the intensities of corresponding regions, and achieve a more efficient comparison of a film with previous ones. Even if the temporal mammogram pair is registered, variations in the image intensities due to imaging conditions can't be factored out. Combining the 2 methods (as discussed in chapter 4) can provide us with a significant framework for temporal comparison of h_{int} based breast-tissue quantitative measures.

2 Computation of the hint surface

In order to calculate the thickness of interesting tissue between each pixel and the X-ray source, we first need to calculate the primary radiation energy $E_p(x,y)$ imparted to the intensifying screen. If the mammogram is digitised inhigh quality, laser scanning devices, the relationship between pixel and density values is linear:

$$P(x,y) = m \cdot D(x,y) + c \qquad (1)$$

After calculating the density image, the total energy imparted to the intensifying screen can be obtained by (2):

$$D(x,y) = \gamma \log_{10}(\beta E_{pse}^{imp}(x,y)) \qquad (2)$$

Where β is a linear constant related the film-speed, and γ is the film screen gradient and can be calculated from the film-screen characteristic curve as described in [*, Chapter 2]. After calculating the total energy imparted in the intensifying screen the primary component is given by (3):

$$E_p^{imp}(x,y) = E_{pse}^{imp}(x,y) - E_s^{imp}(x,y) - E_e^{imp}(x,y) \qquad (3)$$

Where Es(x,y) and Ee(x,y) are the scattered and extra-focal radiation that are modelled and calculated in ([4], Chapter 3) and ([4], Chapter 4) respectively. The last step is to combine equations (4) and (5) to calculate $h_{int}(x, y)$ for all the image pixels:

$$h\mu(E,x,y) = h_{\text{int}}(x,y)\mu_{\text{int}}(E) + h_{fat}(x,y)\mu_{fat}(E) = h_{\text{int}}(x,y)(\mu_{\text{int}}(E) - \mu_{fat}(E)) + H\mu_{fat}(E) \qquad (4)$$

$$E_p(x,y) = \phi(V_{tube},x,y)A_p t_s \int_0^{V_{tube}} N_0^{rel}(E)ES(E)G(E)\cdot e^{-\mu_{luc}(E)h_{plate}} e^{-h\mu(E,x,y)}dE \qquad (5)$$

Where ϕ is the photon flux corresponding to an X-ray tube voltage of V_{tube}, A_p is the pixel area, t_s is the time of exposure, $N_0{}^{rel}(E)$ is the relative number of photons at energy E, $S(E)$ is the absorption ratio of the screen to primary photons of energy E, $G(E)$ is the transmission ratio of the grid for primary photons of energy E, $\mu_{luc}(E)$ is the linear attenuation coefficient of lucite at E and h_{plate} is the thickness of the lucite compression plate. In the system of equations of (4) and (5), the only unknown is $h_{int}(x, y)$.

By exploiting the hint representation, it is possible to simulate a scatter-free monoenergetic simulation thus enhancing the image [4]. Consequently, it is possible to reduce the radiation dose to the patient by removing the anti-scatter grid and enhancing the image using the representation [101].

The h_{int} representation of a temporal mammogram pair can be represented by:

$$\begin{pmatrix} D^{t_1}(x,y) \\ D^{t_2}(x,y) \end{pmatrix} \xrightarrow{eq.(2)\ to(5)} \begin{pmatrix} h_{int}^{t_1}(x,y) \\ h_{int}^{t_2}(x,y) \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} D_{enhanced}^{t_1}(x,y) \\ D_{enhanced}^{t_2}(x,y) \end{pmatrix} \qquad (6)$$

Thus we can use this representation for registration or matching of temporal mammograms either on the h_{int} domain or in the enhanced mammogram pair. These ideas are summarised in (7):

$$\begin{pmatrix} D^{t_1}(x,y) \\ D^{t_2}(x,y) \end{pmatrix} \xrightarrow{eq.(2)\ to(5)} \begin{pmatrix} h_{int}^{t_1}(x,y) \\ h_{int}^{t_2}(x,y) \end{pmatrix} \xrightarrow{T(x,y)} \begin{pmatrix} T[h_{int}^{t_1}(x,y)] \\ h_{int}^{t_2}(x,y) \end{pmatrix} \xrightarrow{(2)} \begin{pmatrix} D_{aligned}^{t_1}(x,y) \\ D_{aligned}^{t_2}(x,y) \end{pmatrix} \qquad (7)$$

Where T(x, y) is the non-rigid transformation (thin-plate spline interpolating function in our case), that approximates the difference in compression between the temporal pair. In chapter 5, a method for preserving the h_{int} image information after registration was suggested.

Appendix D: Basic concepts related to diagnostic tests

In this Appendix, some of the more common concepts related to Information Mastery and Evidence Based Medicine (EBM) for clinical diagnosis are described. A more analytical description can be found in [192].

Clinical information from patients is routinely collected in all hospital across the country. However, in order to understand the correlation of the different clinical signs, symptoms and diagnostic tests to the likelihood of a disease, this information has to be accurately analysed and interpreted through objective measures and definitions.

The following definitions are related to **clinical diagnostic tests** and the main interest from the image analysis standpoint is to improve the understanding of clinical articles as well as the reported results on CAD systems:

Sensitivity: Is the probability of a positive test among patients with a disease. For example if a diagnostic test was positive in 85 out of 100 patients previously diagnosed with breast cancer (in 15 cases the test was negative), then the sensitivity of the test is 85%.

Specificity: It is the probability of a test being negative among "healthy" patients. Again, the number of negative tests (among patients without the disease) has to be divided with the total number of tests to calculate the specificity.

Incidence: Is the probability that a healthy patient develops the disease during an interval (e.g. in a year, or in the screening interval for the case of breast cancer). This definition is important when determining the incidence of breast cancer, since a significant number of cancers are missed. In chapter 4 we suggested the use of image registration to determine (retrospectively) the number of "true" interval cancers.

Prevalence: Is the probability of a disease in the entire population. Using the definitions of sensitivity, specificity and prevalence the probability of disease given a positive test can be defined according to Bayes'equation:

$$P(\text{disease} \mid \text{positive test}) = \frac{\text{prevalence} \cdot \text{sensitivity}}{[\,(\text{prevalence} \cdot \text{sensitivity}) + ((1 - \text{prevalence}) \cdot (1 - \text{sensitivity}))\,]}$$

Although, sensitivity and specificity are important in describing diagnostic tests, they do not give plenty of information to interpret the results of a test. For this reason, the predictive values are more useful to the clinicians:

Positive predictive value: Is the probability of disease among patients that had a positive test.

Negative predictive value: Is the probability of no disease among patients that has a negative test.

For example, if we want to describe a diagnostic test foe breast cancer (e.g. a CAD systems that detects malignant masses) and we get the following results:

	Patients with breast cancer	Patient without breast cancer
Positive test	89 (true positives)	40 (false positives)
Negative test	11 (false negatives)	60 (true negatives)

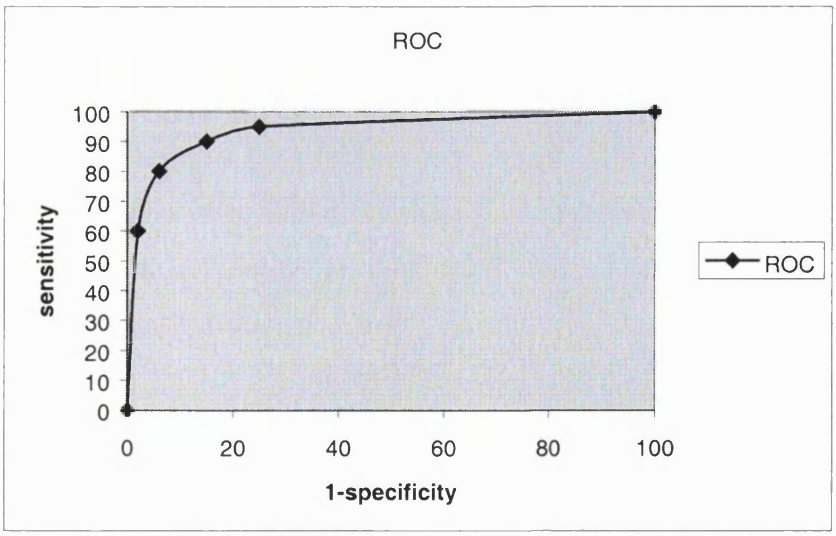
Then we can calculate the following:
Sensitivity = 89 / 100 = 89%, Specificity = 40 /100= 40%
Positive predictive value = 89 / (89 + 40) = 69%
Negative predictive value = 11 / (11 + 60) = 15%

This means that although the sensitivity of the test is very high the probability of disease among patients with a positive test is 69%, which is the probability that a patient has breast cancer if the CAD test for cancer is positive.
If now, we can redefine what is "positive" or "negative" according to the CAD (e.g. change the threshold at which a detected mass is classified as abnormal), the values for the specificity and sensitivity will change. By doing that for several thresholds we can

plot the values of sensitivity vs. (1-specificity). This is called a receiver-operating characteristic curve (ROC). Examining different thresholds represent the trade-off between sensitivity and specificity, or between false positives and negatives. The area under the ROC curve is measure of good the test is in discriminating between "healthy" and non-"healthy" patients and can be used for comparison with other tests. Following the previous example, lets assume that by changing the threshold at which (e.g. based on texture) a mass was classified as malignant, we got the following values for the specificity and sensitivity:

Threshold that defines malignancy	Sensitivity	Specificity
Threshold 1	60%	98%
Threshold 2	80%	94%
Threshold 3	90%	85%
Threshold 4	95%	75%

Then the ROC curve shows that the test is adequate for discriminating between benign and malignant masses:



Appendix E: Gallery of Results

Registration of bilateral pairs

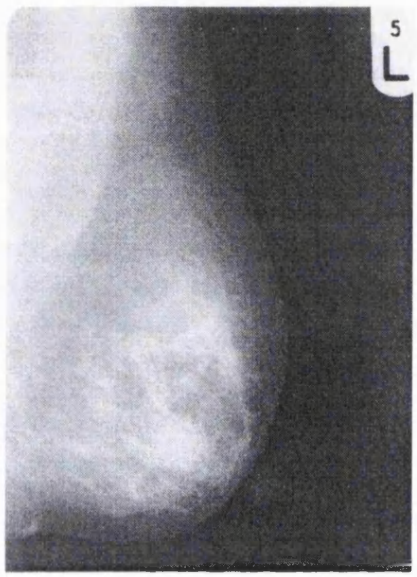
In this section, we present results of bilateral mammograms. In each case the source image is registered to the target and the resulting, transformed image is shown together with the transformed grid and the difference image after registration. In some cases, we also show the subtraction image before registration and the joint histograms (before and after registration).

In the end of each case we include some comments and in certain cases some clinical information that is derived from the registration of the image pair. The following cases are presented in this section:

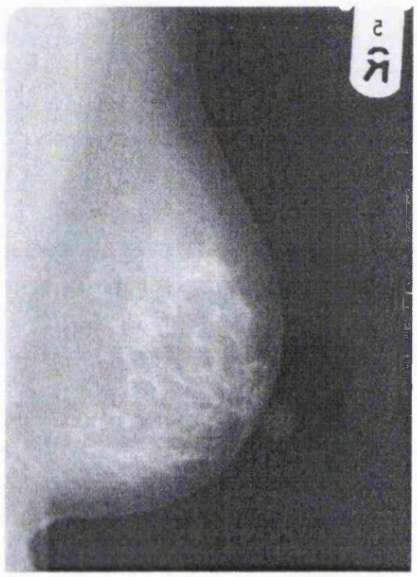
Case 1: Bilateral mammogram pair of a post-menopausal woman. There is a mass in the left breast that is highlighted after registration.
Case 2: Bilateral mammogram pair of a post-menopausal woman. The right mammogram is slightly denser in the nipple plane.
Case 3: Bilateral mammogram pair of a post-menopausal woman. Registration highlights a mass in the lower quadrants of the right breast.
Case 4: Bilateral mammogram pair of a post-menopausal woman. A cancer in the left breast is pronounced after registering the images.
Case 5: Bilateral mammogram pair of a post-menopausal woman. An increased density in the right breast and a "suspicious" density in the left breast near the axilla are highlighted after registration.
Case 6: Bilateral mammogram pair of a post-menopausal woman. A small mass is present in the lower quadrants of the right breast

The registration results on these cases are now presented in detail:

CASE 1:



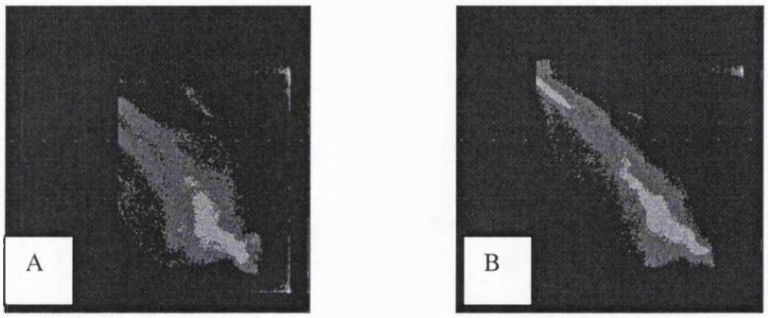
TARGET



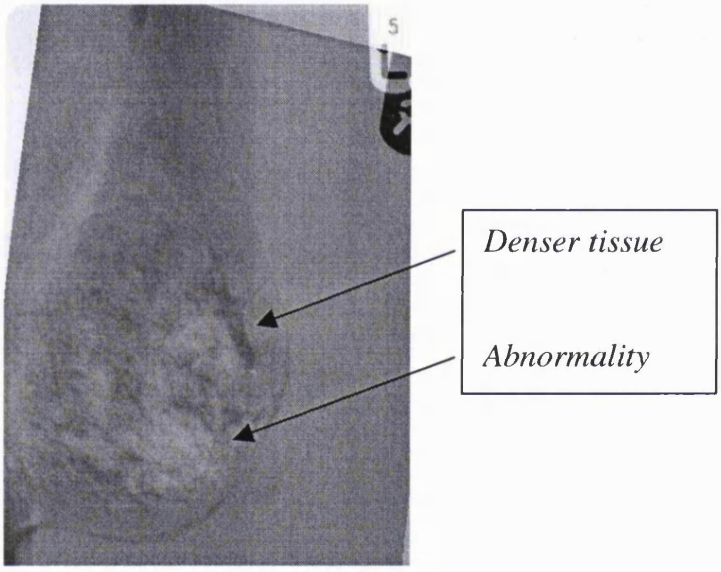
SOURCE

TARGET

SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION

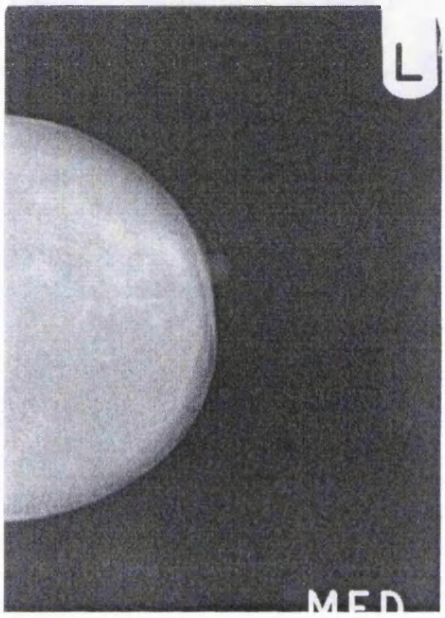


DIFFERENCE IMAGE AFTER REGISTRATION

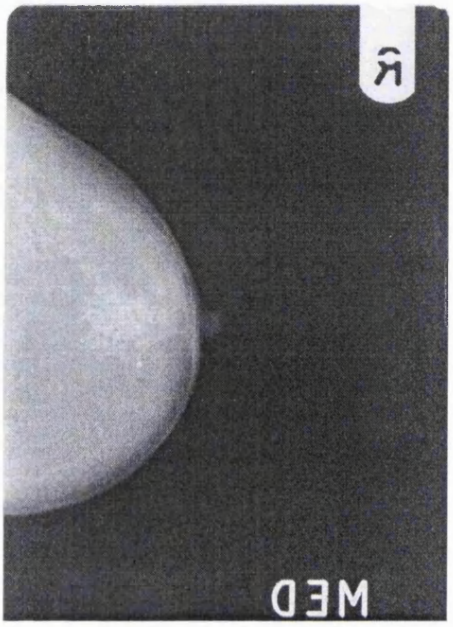
COMMENTS

There is a cancer present in the left breast. Image registration allows an automated comparison of the differences in tissue structure between the bilateral pair. As shown in the difference image the two areas of "significant " differences correspond to a denser tissue region and an abnormality present in the left (target) mammogram.

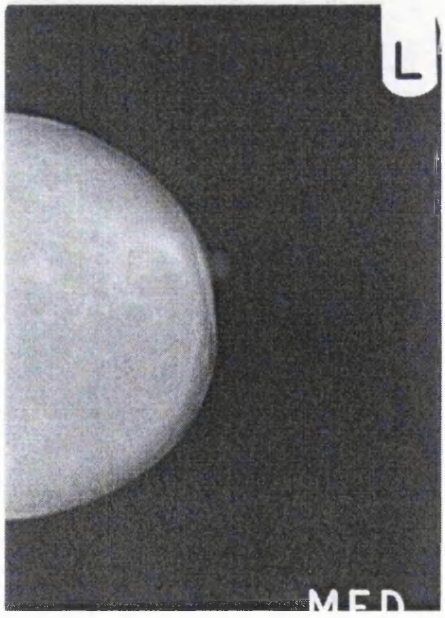
CASE 2:



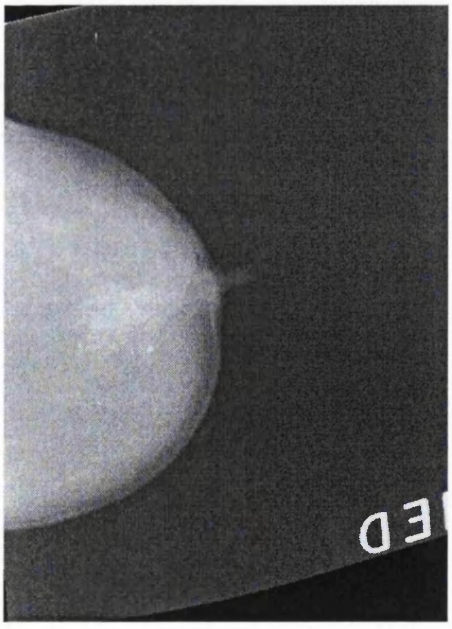
TARGET



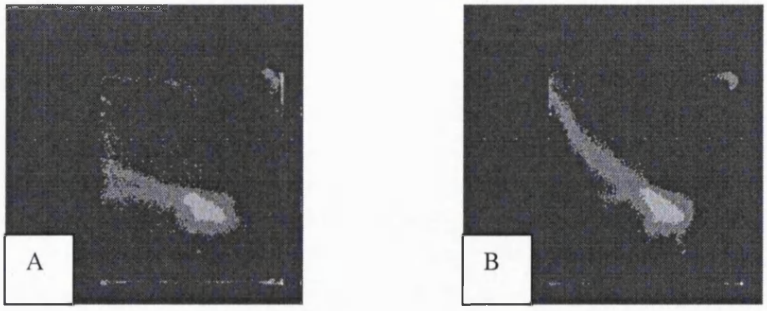
SOURCE



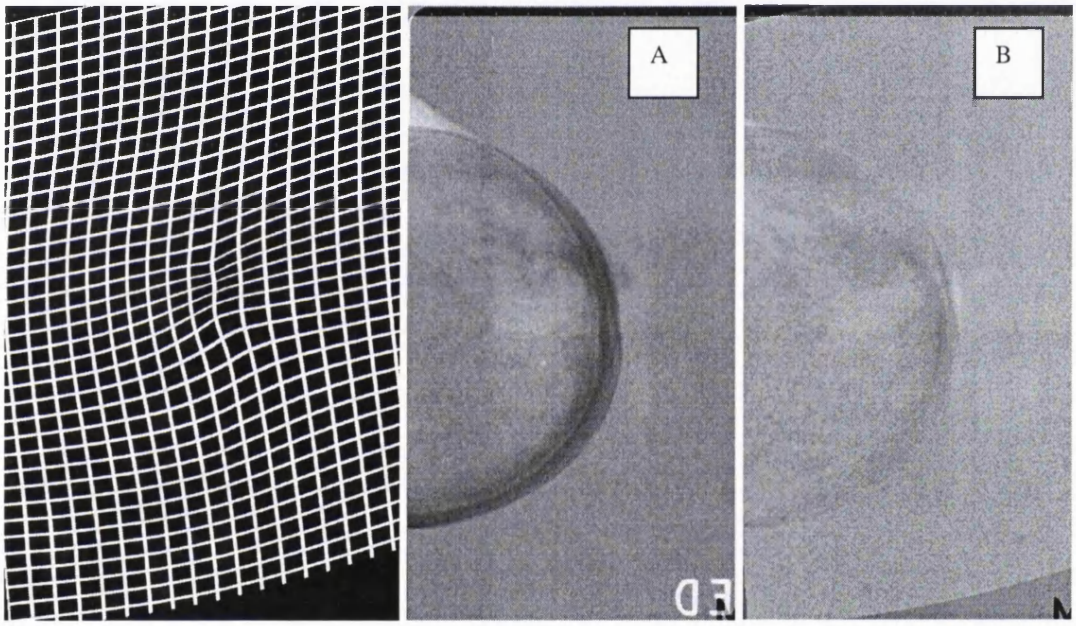
TARGET



SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION

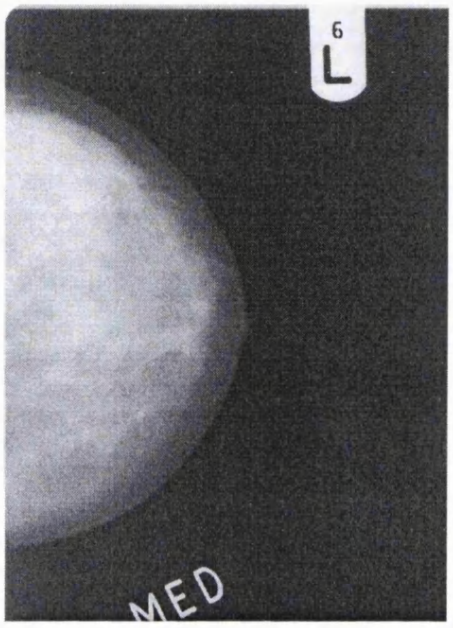


THE TRANSFORMATION GRID AND THE DIFFERENCE IMAGE BEFORE (A) AND AFTER (B) REGISTRATION

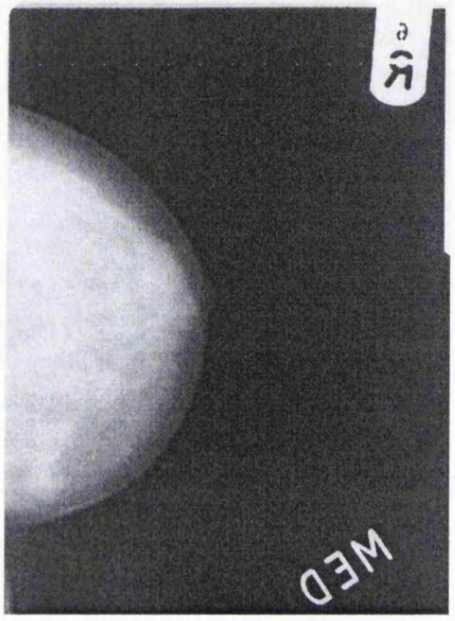
COMMENTS

In this pair of bilateral mammograms there is a difference in the tissue architecture near the nipple. As shown in the joint histogram images after the images are aligned the intensity pairs are more concentrated in the main diagonal. The good overlap is also confirmed by the difference image after registration

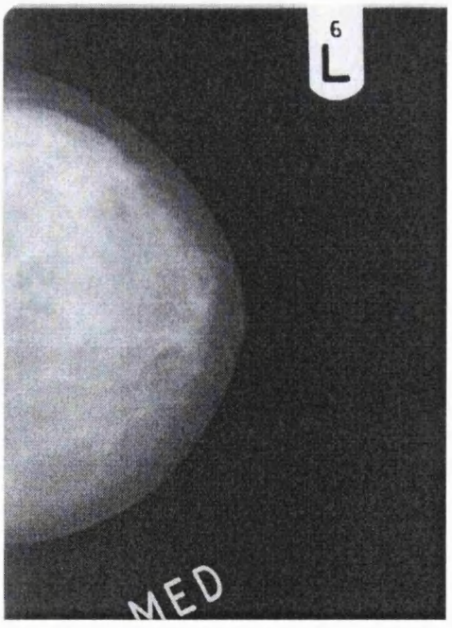
CASE 3:



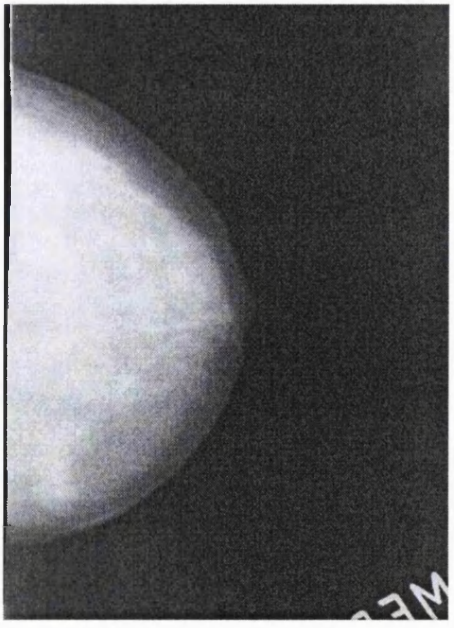
TARGET



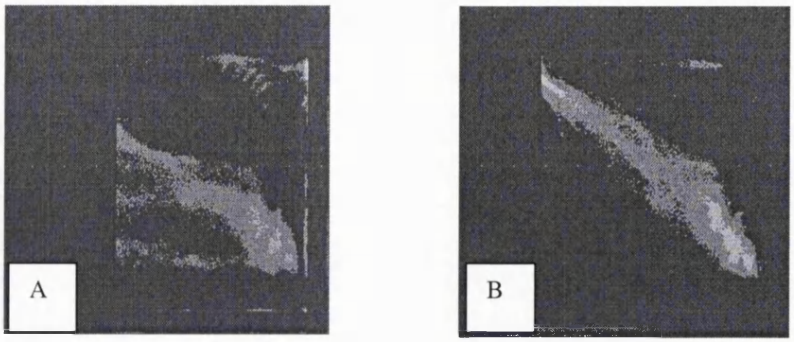
SOURCE



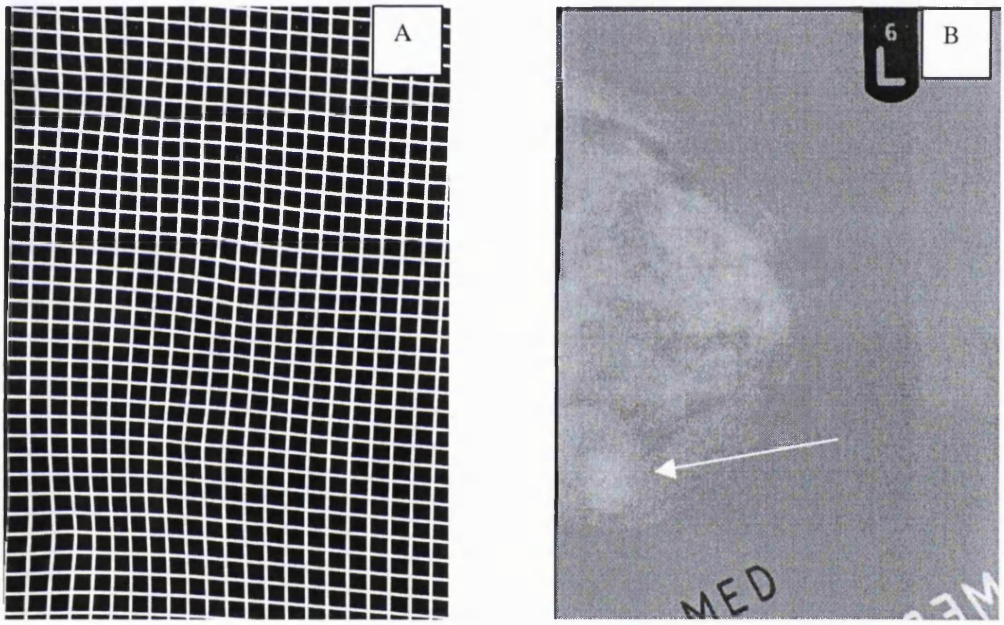
TARGET



SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION

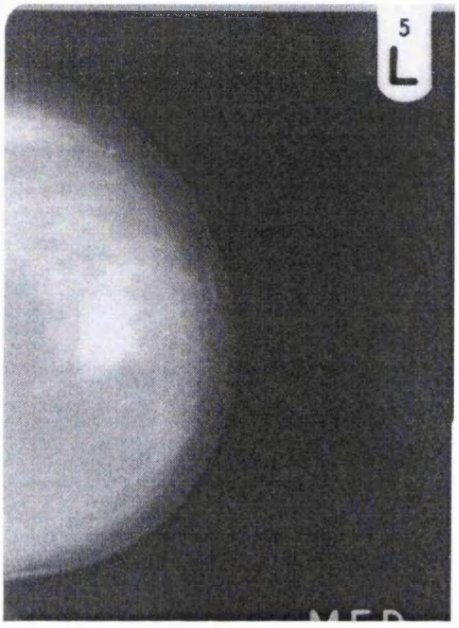


THE TRANSFORMATION GRID (A) AND THE DIFFERENCE AFTER (B) REGISTRATION

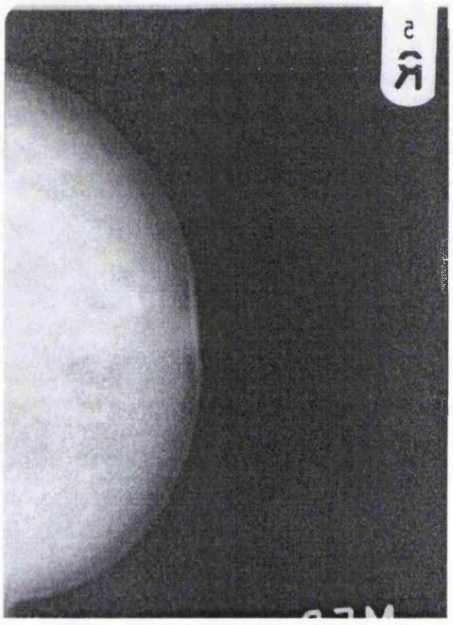
COMMENTS

In this pair of bilateral mammograms the right breast appears to be denser than the left. As shown in the joint histogram images after the images are aligned the intensity pairs are more concentrated in the main diagonal. In the difference image after registration the region that exhibits the larger intensity difference (shown by the arrow) corresponds to the mass present in the lower quadrants of the right breast.

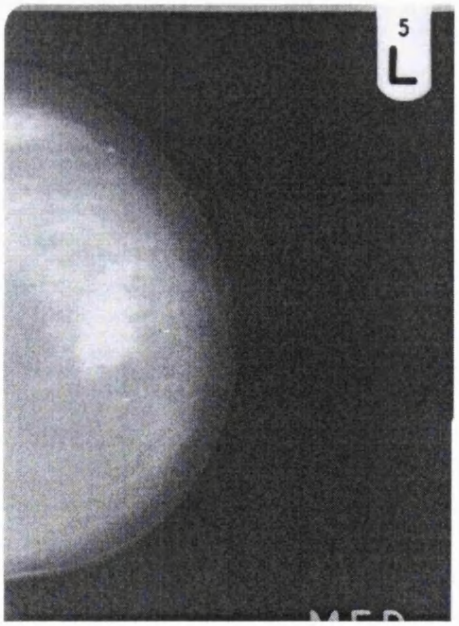
CASE 4:



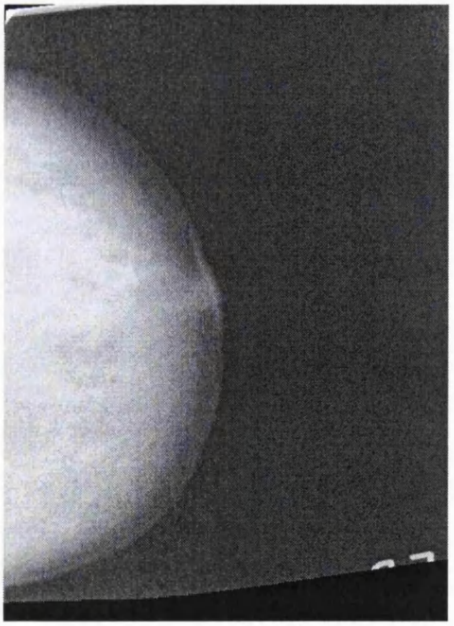
TARGET



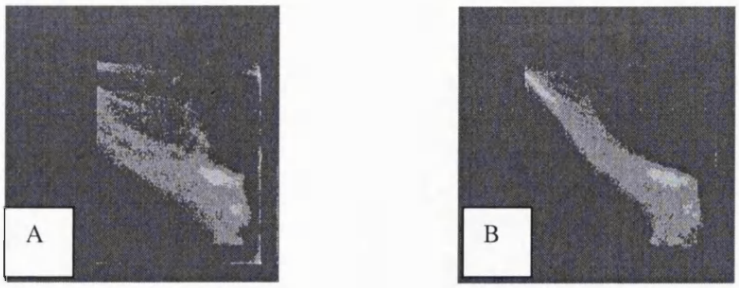
SOURCE



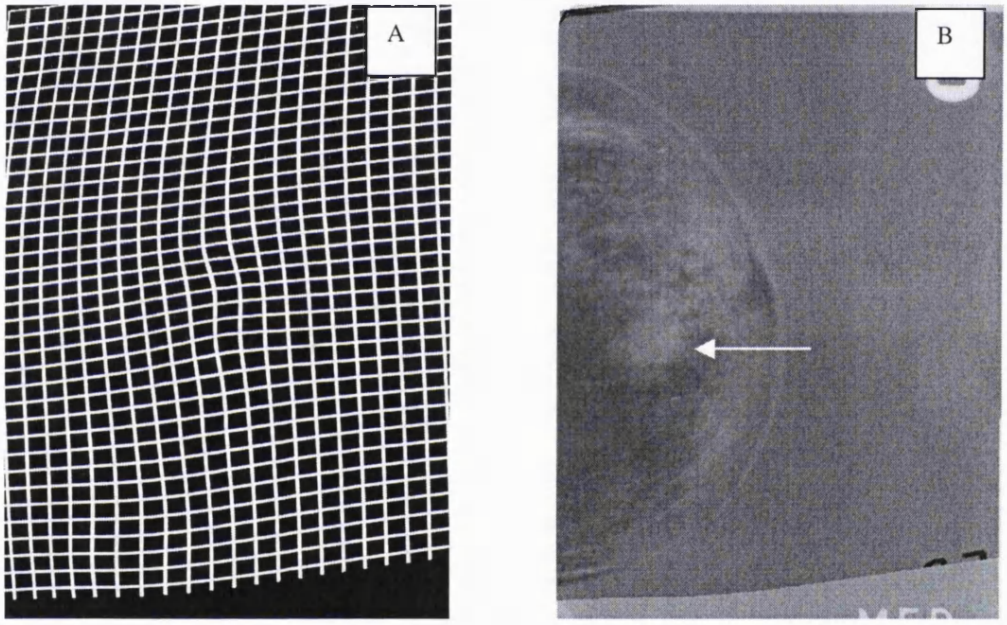
TARGET



SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION



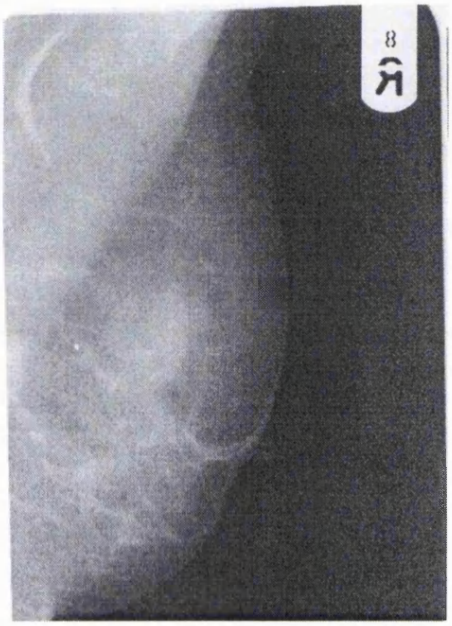
THE TRANSFORMATION GRID (A) AND THE DIFFERENCE AFTER (B) REGISTRATION

COMMENTS

The cancer in the left breast appears as the dominant dissimilarity in the difference image after registration as the rest of the breast is aligned. The arrow shows the "bright" region of the cancer in the difference image.

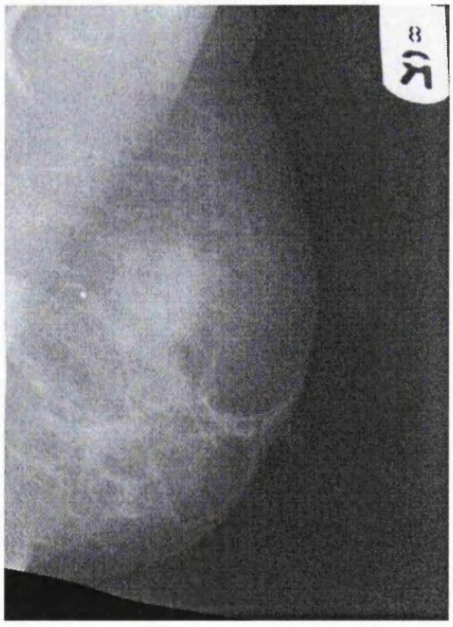
CASE 5:

TARGET

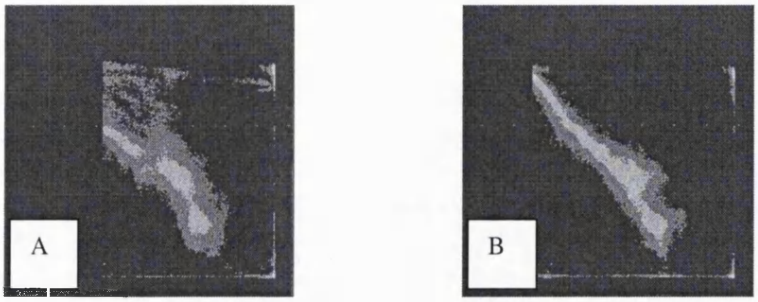


SOURCE

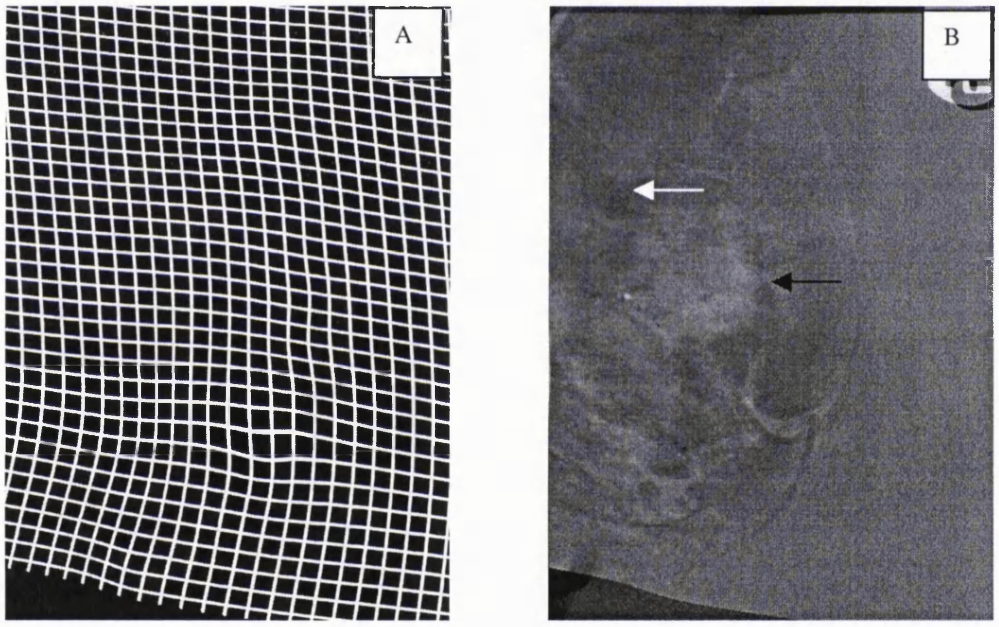
TARGET



SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION

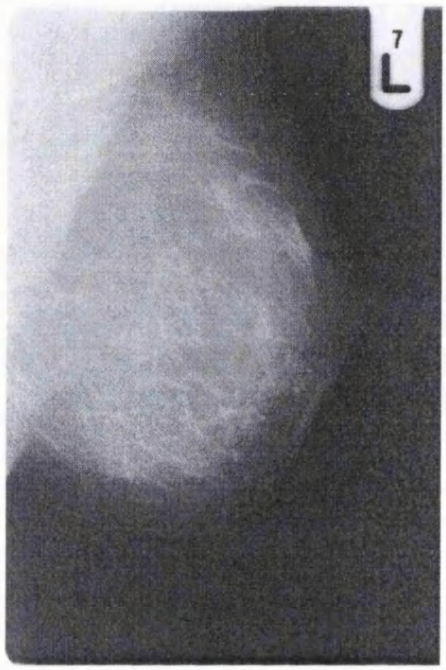


THE TRANSFORMATION GRID (A) AND THE DIFFERENCE AFTER (B) REGISTRATION

COMMENTS

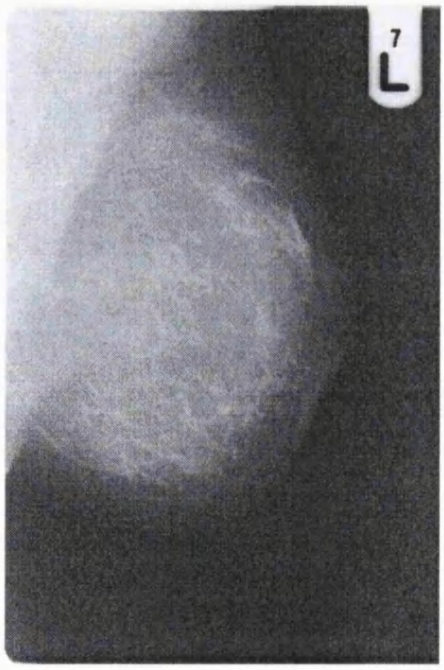
In this bilateral pair there are two different between the left and right mammogram. As is highlighted in the difference image after registration, there is a dense region present only in the right mammogram (indicated by the black arrow), and a small "suspicious" density near the axilla (indicated by the '"white" arrow) in the left breast.

CASE 6:

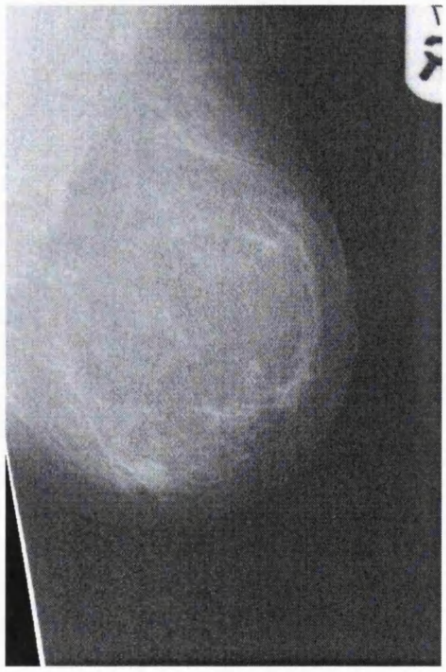


TARGET

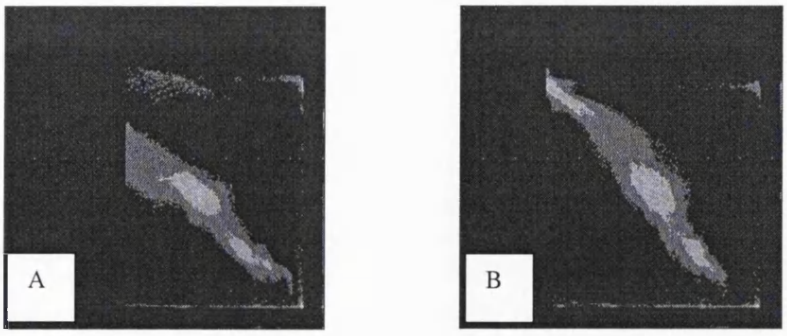
SOURCE



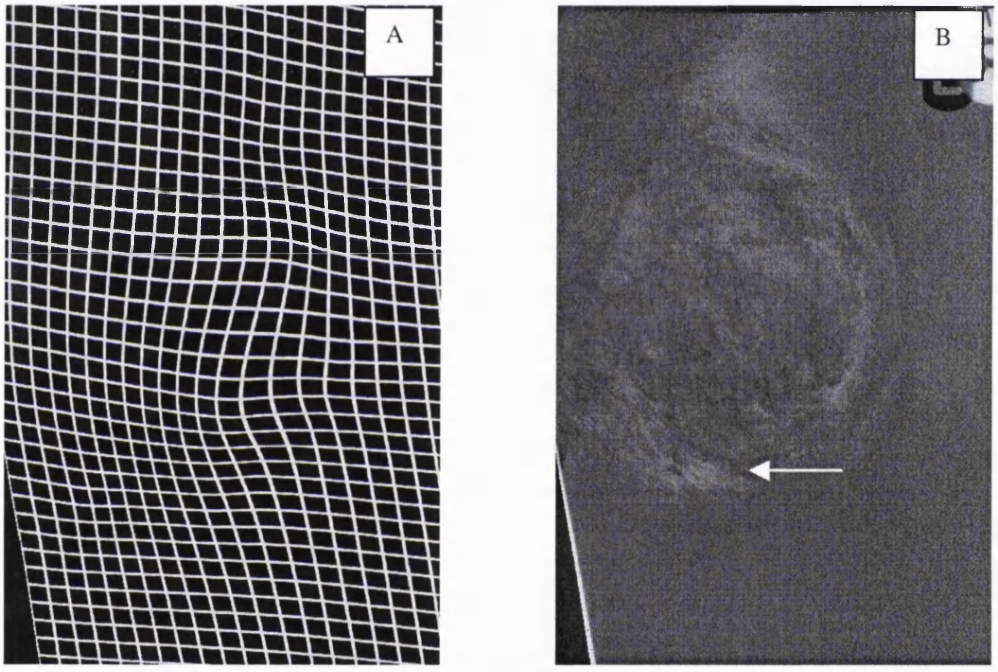
TARGET



SOURCE TRANSFORMED



JOINT HISTOGRAM BEFORE (A) AND AFTER (B) REGISTRATION



THE TRANSFORMATION GRID (A) AND THE DIFFERENCE AFTER (B) REGISTRATION

COMMENTS

After registering the bilateral air, the main difference in the subtraction image is a small mass in the lower quadrants of the right breast (indicated by the arrow in the difference image).

Registration of temporal HRT pairs

In this section, we present results of temporal mammograms. In each case the source image is registered to the target and the resulting, transformed image is shown together with the transformed grid and the difference image after registration. In some cases, we also show the subtraction image before registration and the joint histograms (before and after registration).

In the end of each case we include some comments and in certain cases some clinical information that is derived from the registration of the image pair. The following cases are presented in this section:

Case 1: HRT user (5 years). The mammogram where the cancer was diagnosed is compared with a previous one.

Case 2: HRT user (2 years). HRT was stopped because a cancer was detected in the right breast. A temporal sequence of the left breast is registered highlighting a new mass.

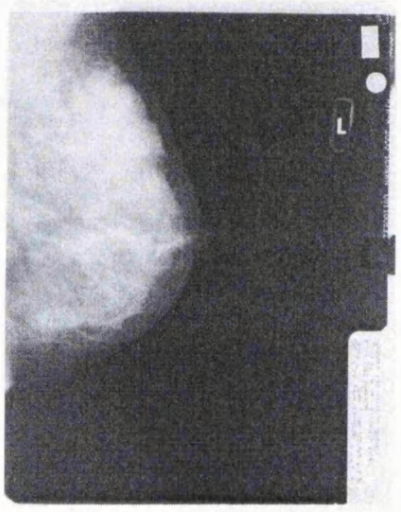
Case 3: HRT user (continuing). A temporal mammogram sequence is registered in order to highlight the regions of density regeneration.

Case 4: HRT user due to hysterectomy. Four mammograms after HRT cessation are registered.

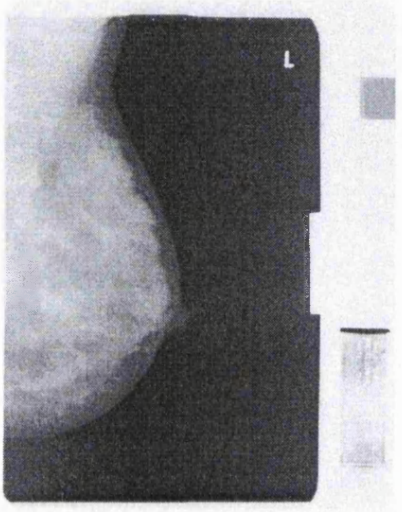
Case 5: Five mammograms (two before and two after HRT cessation are registered) highlighting the changes in breast tissue density (first increase and then decrease).

The registration results on these cases are now presented in detail:

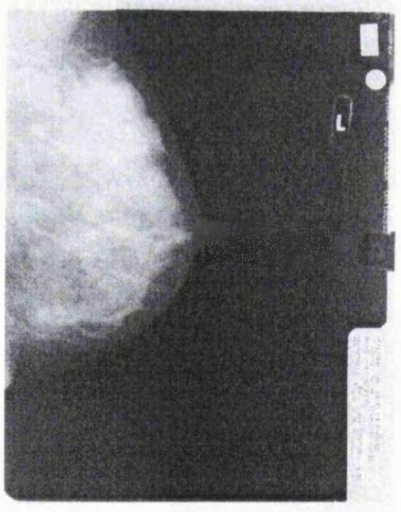
CASE 1:



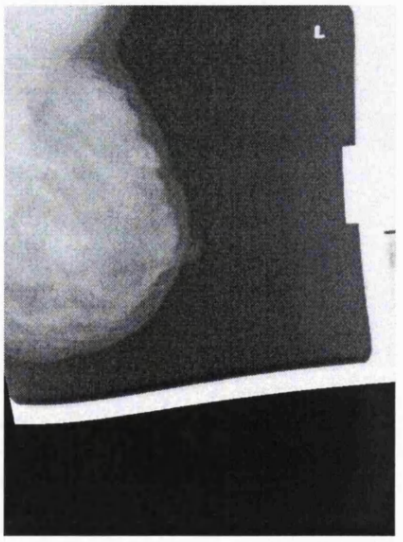
TARGET (L ML 98)



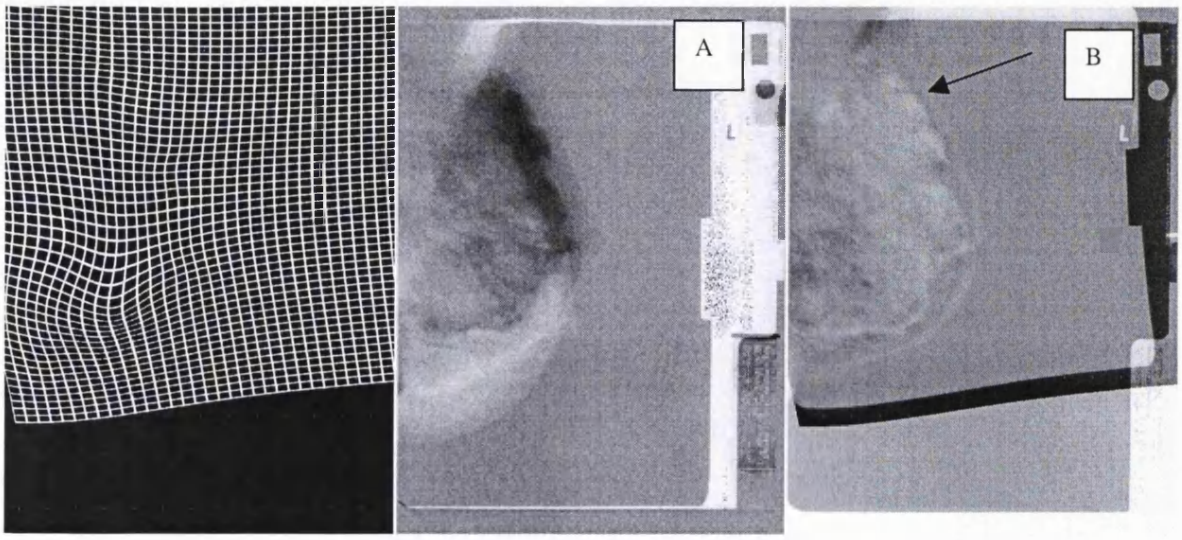
SOURCE (L ML 96)



TARGET



SOURCE TRANSFORMED

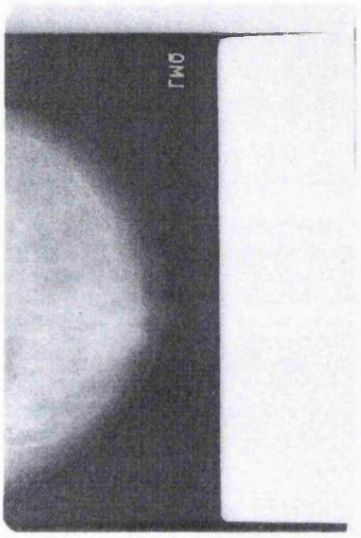


THE TRANSFORMATION GRID AND THE DIFFERENCE IMAGE BEFORE (A) AND AFTER (B) REGISTRATION

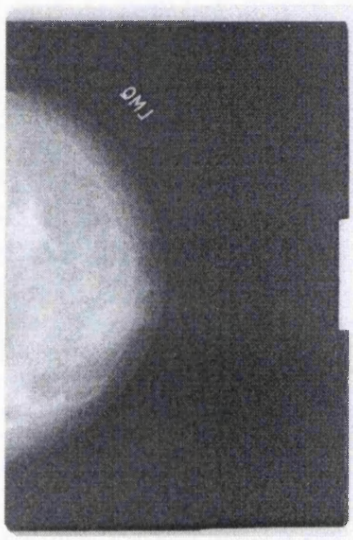
COMMENTS

The patient was born in 1931 and had her menopause in 1984. The use of HRT started in 1993 until 1998 when a cancer was diagnosed in the left breast. Here, we register the left ML mammograms from 1996 and 1998 (when the cancer was diagnosed). After registration there is a small number of "significant difference" regions in the difference image. The arrow indicates the region that corresponds to the cancer that was diagnosed in the left ML mammogram (1998).

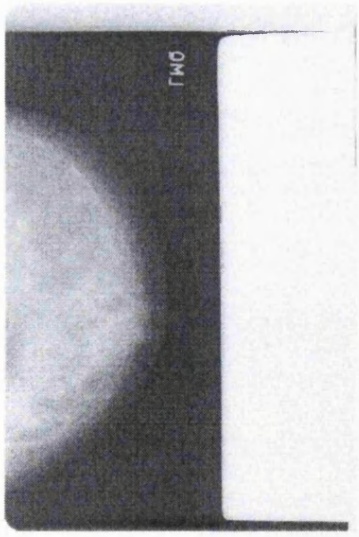
CASE 2:



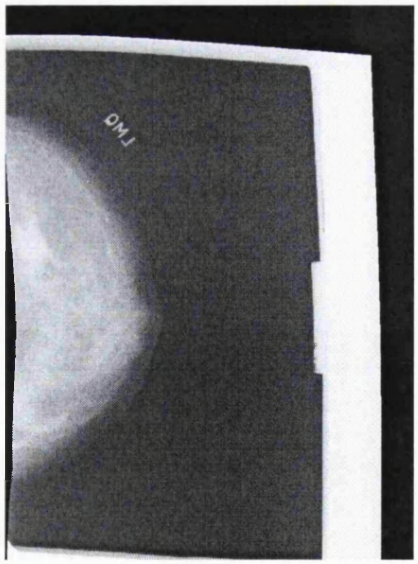
TARGET (L LMQ 95)



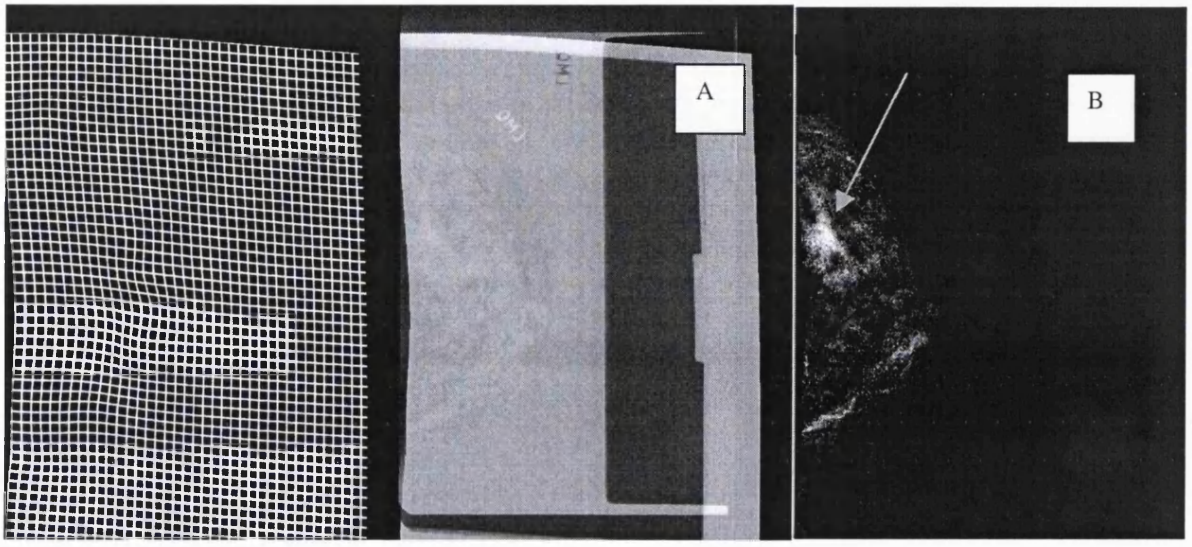
SOURCE (L LMQ 96)



TARGET



SOURCE TRANSFORMED

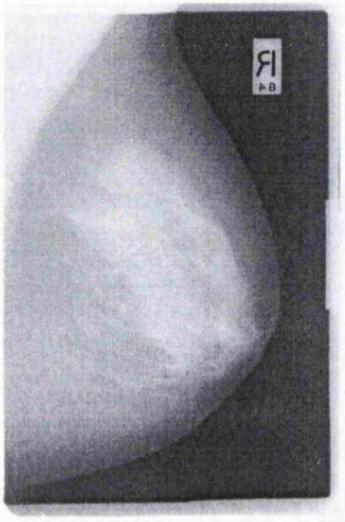


THE TRANSFORMATION GRID AND THE DIFFERENCE IMAGE AFTER REGISTRATION (not processed (A), thresholded (B))

COMMENTS

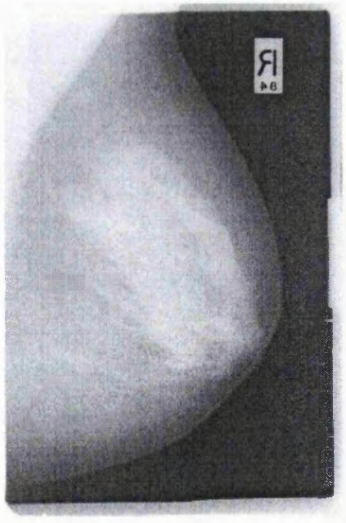
The patient was born in 1944 and had her menopause in 1991. The use of HRT started in 1992 until 1994 when a cancer was diagnosed in the right breast. Here, we register the left LMQ mammograms from 1995 and 1996. After registration there is one predominant region of significant difference in the subtraction image (indicated by the arrow). That corresponds to a mass developed in the left breast (L LMQ 96).

CASE 3:

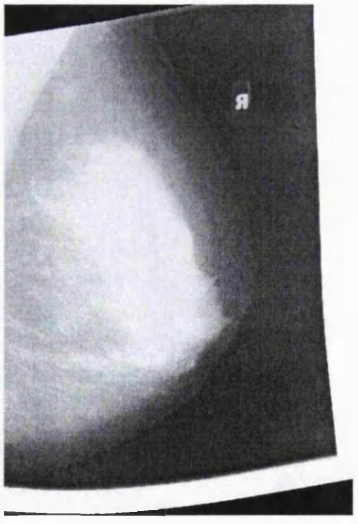


TARGET (R ML 97)

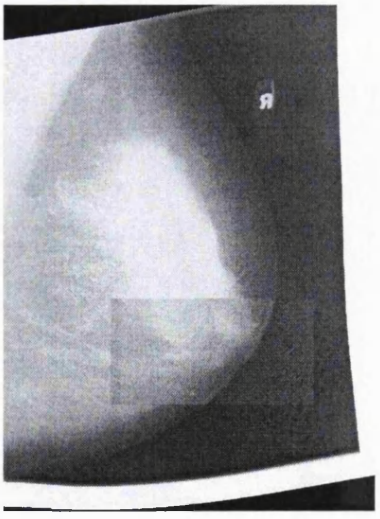
SOURCE (R ML 98)



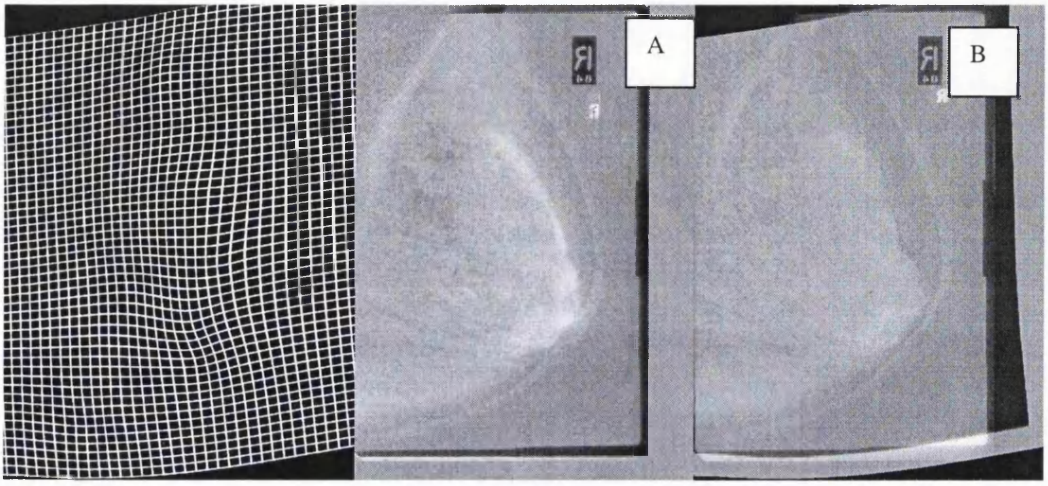
TARGET



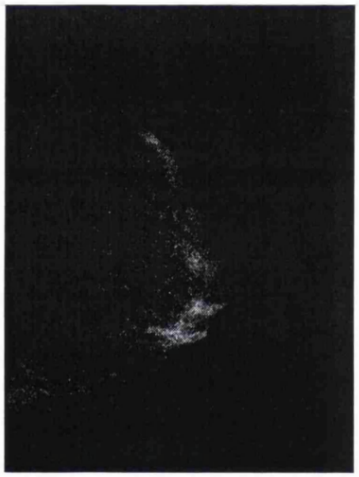
SOURCE TRANSFORMED



MIXED



THE TRANSFORMATION GRID AND THE DIFFERENCE IMAGE BEFORE (A) AND AFTER (B) REGISTRATION

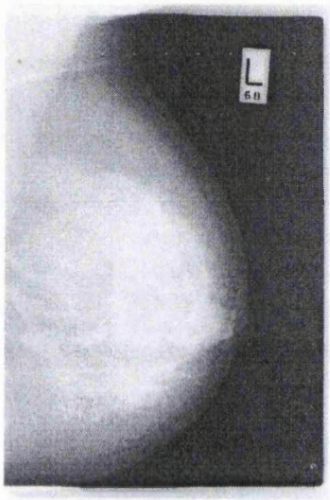


DIFFERENCE IMAGE AFTER REGISTRATION (THRESHOLDED)

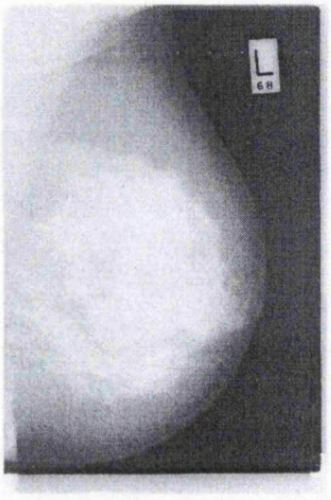
COMMENTS

The patient was born in 1937 and had her menopause in 1992. The use of HRT started in 1992. Here, we register the right ML mammograms from 1997 and 1998. The "MIXED" image after registration, shown in the previous page, could assist the clinician to better understand the local variations in tissue density over time, and their possible implications. Obviously, in this case there is an overall increase in density due to HRT. From the difference image we see that the most significant area of change is near the nipple.

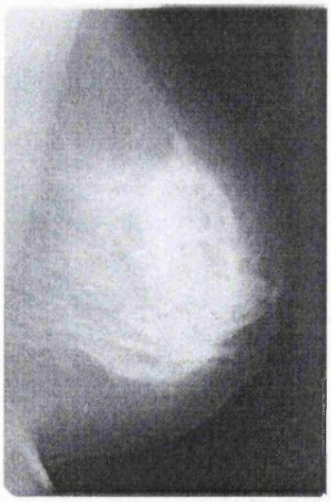
CASE 4:



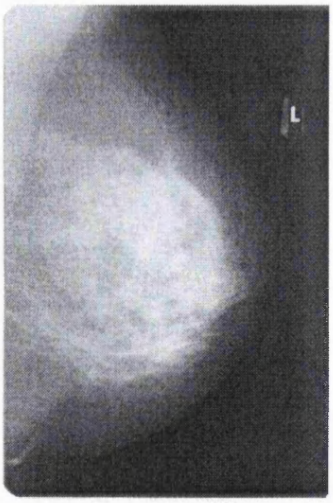
TARGET (L ML 96)



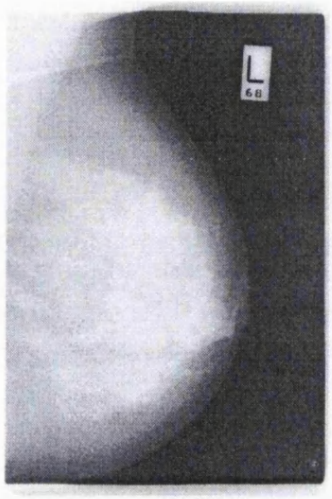
SOURCE 1(L ML 97)



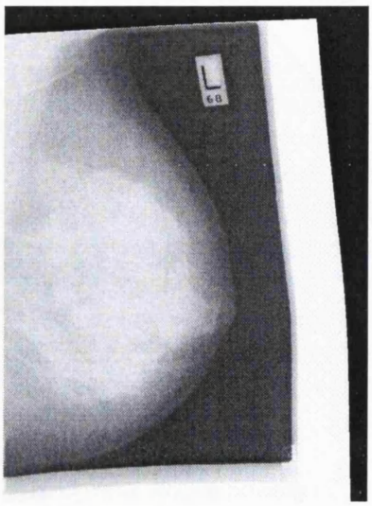
SOURCE 2(L ML 98)



SOURCE 3(L ML 99)

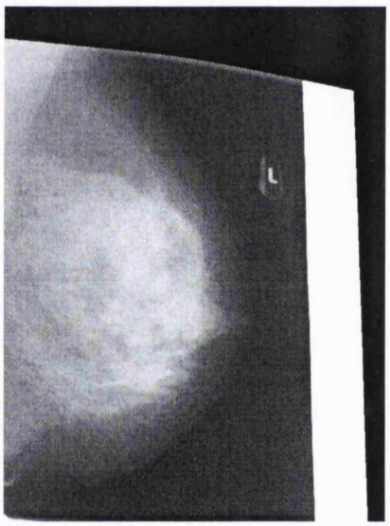


TARGET (L ML 96)



REGISTERED (L ML 97)

REGISTERED (L ML 98)



REGISTEREDL ML 99)

COMMENTS

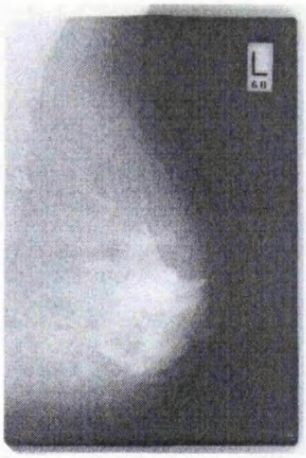
The patient was born in 1932 and had her menopause in 1972 due to hysterectomy. The use of HRT stopped in 1996. Here we register 3 subsequent mammograms to the first one (left ML 1996), and observe the gradual density decrease (due to HRT cessation).

CASE 5:

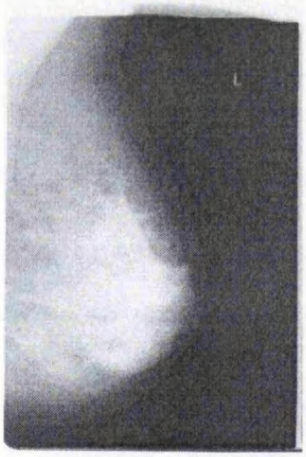
1991

1992(TARGET)

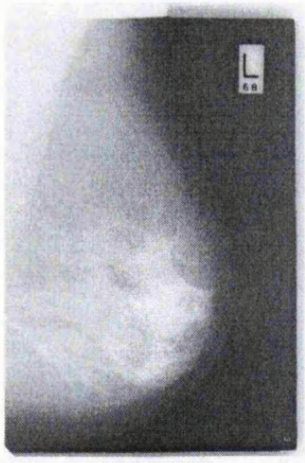
1993



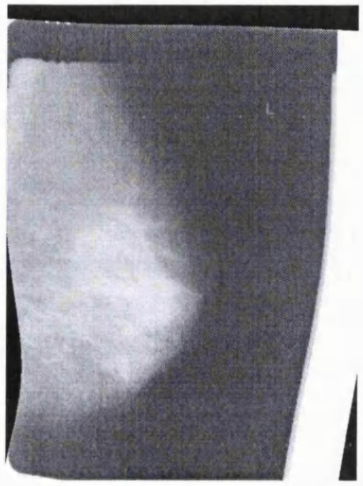
1994



1992(TARGET)



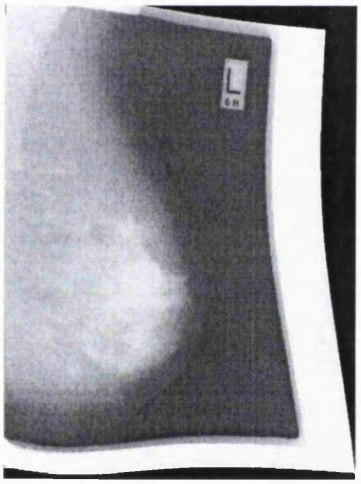
1996



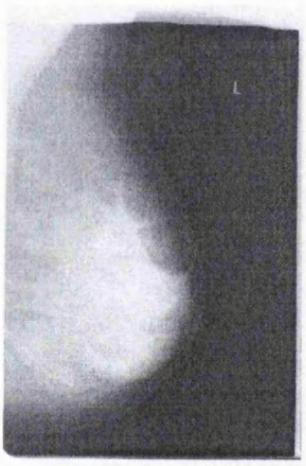
REGISTERED 1991

1992(TARGET)

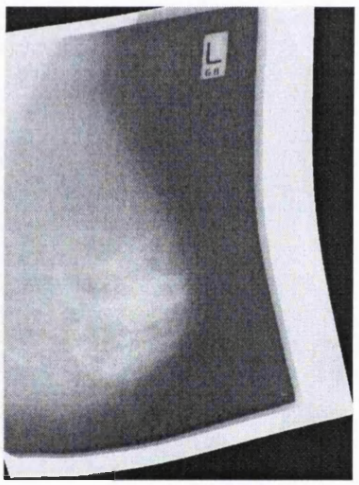
REGISTERED 1993



REGISTERED 1994



1992(TARGET)



REGISTERED 1996

COMMENTS

The patient was born in 1940 and had her menopause in 1970 due to hysterectomy. The use of HRT stopped in 1992. Here we register mammograms before and after 1992 to the left ML (1992). Registration can facilitate the comparison of the images, in this case first the density increase (from 91 to 93 due to HRT) and subsequently decrease (94 to 96 due to HRT cessation).

Registration of normalised image pairs

In this section, we present results of temporal image pairs that have been normalised using the h_{int} representation of interesting tissue (Appendix C). In each case the source image is registered to the target and the resulting, transformed image is shown together with the transformed grid and the difference image after registration. In some cases, we also show the subtraction image before registration and the joint histograms (before and after registration).

In the end of each case we include some comments and in certain cases some clinical information that is derived from the registration of the image pair. The following cases are presented in this section:

Case 1: Temporal h_{int} mammogram pair. No significant changes.

Case 2: Temporal h_{int} mammogram pair. There is a slight increase in density in the lower quadrants.

Case 3: Temporal h_{int} mammogram pair. The breast if highly involuted and there are no significant changes.

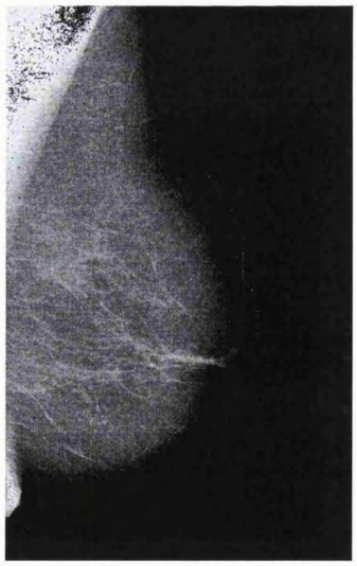
Case 4: Temporal h_{int} mammogram pair. There is a slight decrease in density highlighted by the difference image after registration.

The registration results on these cases are now presented in detail:

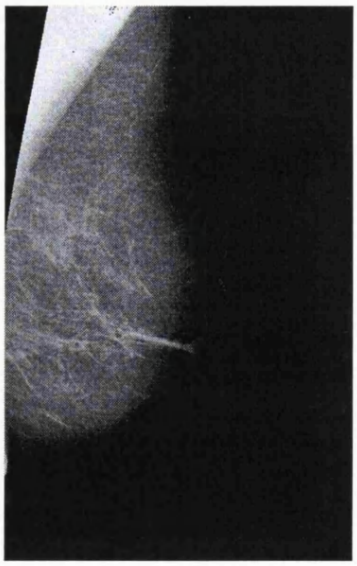
CASE 1:

TARGET

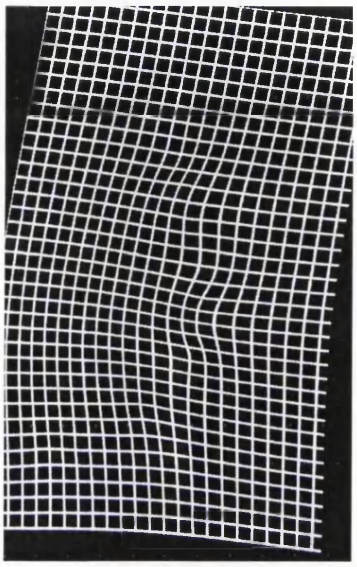
SOURCE



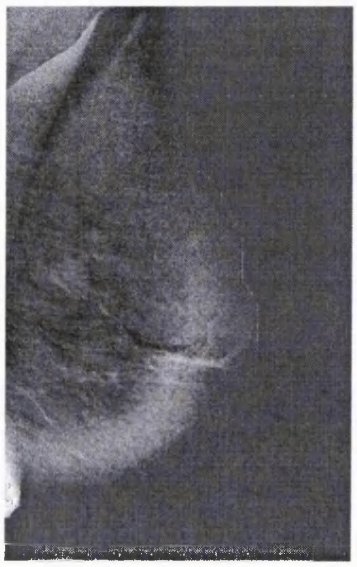
TARGET



SOURCE TRANSFORMED



TRANSFORMED GRID



SUBTRACTION BEFORE

SUBTRACTION AFTER

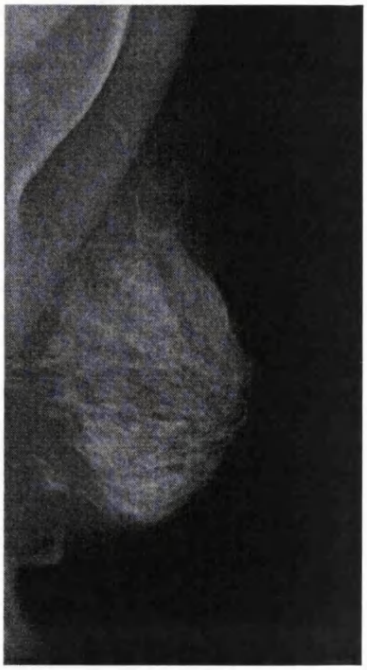
COMMENTS

The image pair is normalised using the h_{int} representation of interesting tissue. There is no significant tissue density change between the mammograms. Due to the difference in compression the images are not aligned, as is highlighted by the difference image before registration (large intensity differences). Most of these differences disappear after aligning the images as is shown in the difference image after registration. The background corresponds to grey-level zero (since the intensities are rescaled in the subtraction image).

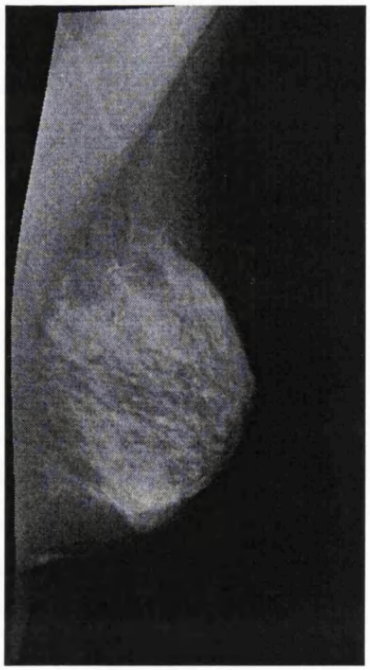
CASE 2:

TARGET

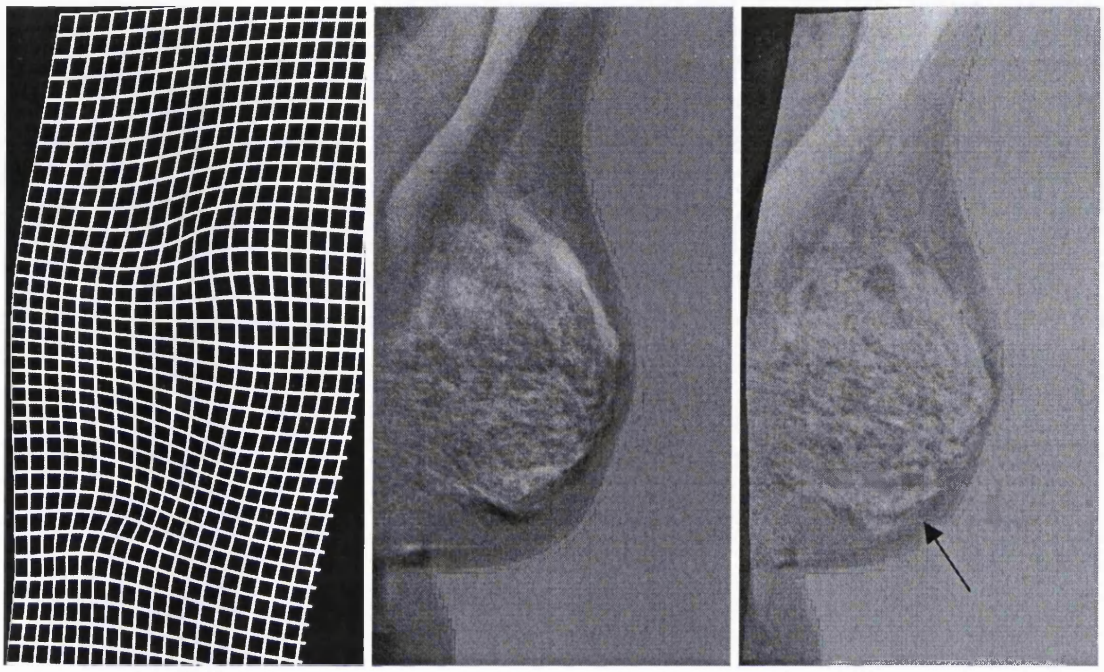
SOURCE



TARGET



SOURCE TRANSFORMED

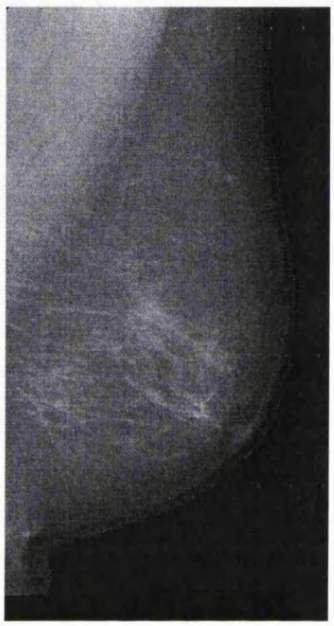


TRANSFORMED GRID SUBTRACTION BEFORE SUBTRACTION AFTER

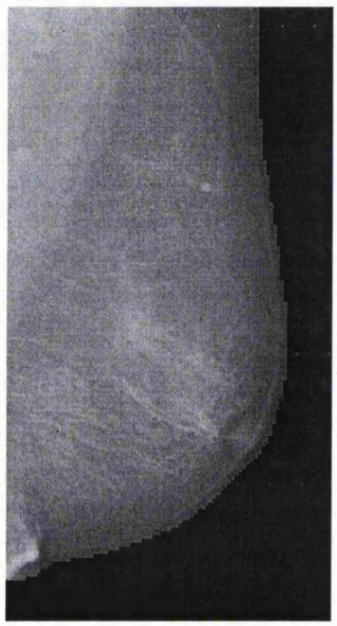
COMMENTS

There is an increase in tissue density in the most recent mammogram. It is obvious that the increase is more dominant in the lower part of the breast. After registration, the subtraction image highlights regions of significant difference. The arrow, in the subtraction image after registration shows the region of increased density in the lower quadrants of the breast (the most recent image is subtracted from the previous one after registration.

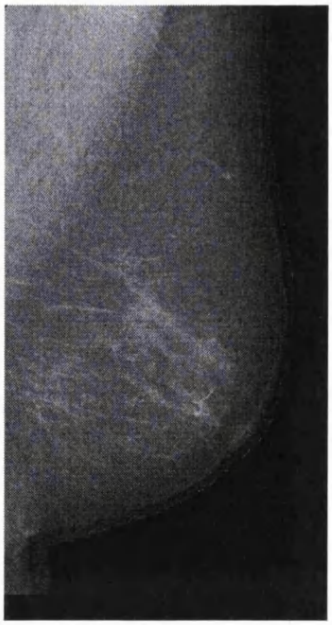
CASE 3:



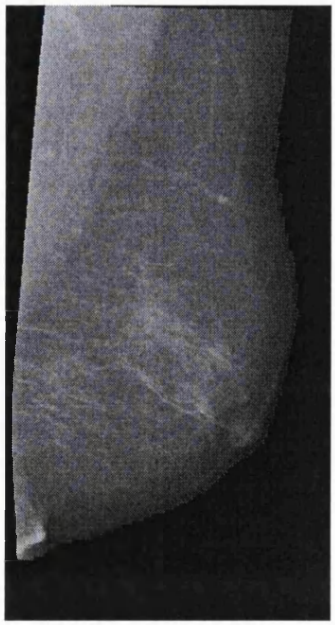
TARGET



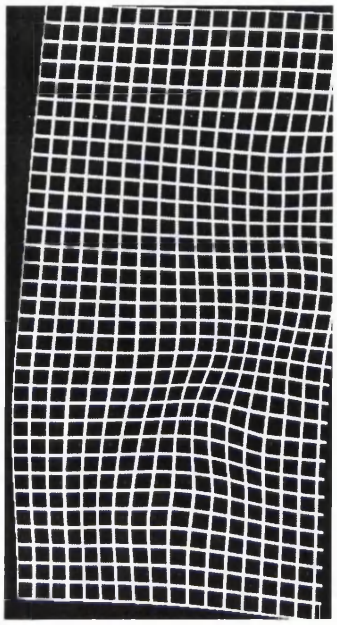
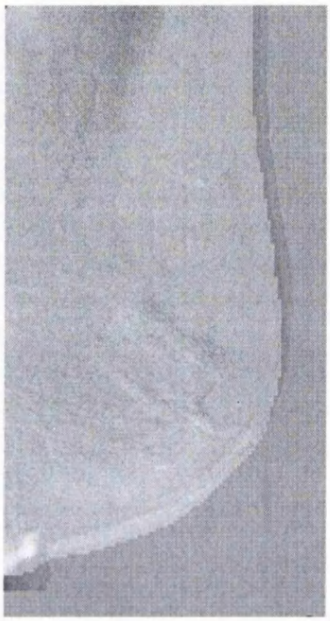
SOURCE



TARGET



SOURCE TRANSFORMED

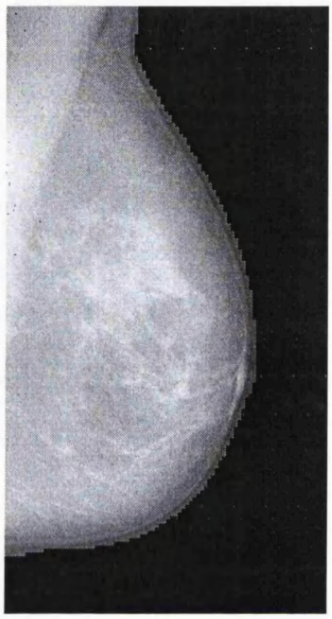

TRANSFORMED GRID SUBTRACTION BEFORE SUBTRACTION AFTER

COMMENTS

The image pair is normalised using the h_{int} representation of interesting tissue. The tissue structure appears to be very similar in the two mammograms. Due to the difference in compression the images are not aligned, as is highlighted by the difference image before registration (large intensity differences) where the glandular structures are misaligned. Most of these differences disappear after aligning the images as is shown in the difference image after registration. The background corresponds to grey-level zero (since the intensities are rescaled in the subtraction image).

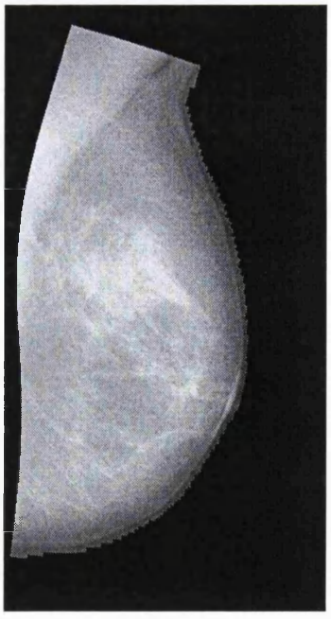
CASE 4:

TARGET

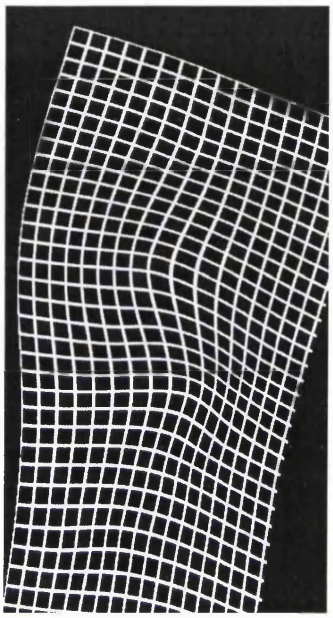


SOURCE

TARGET

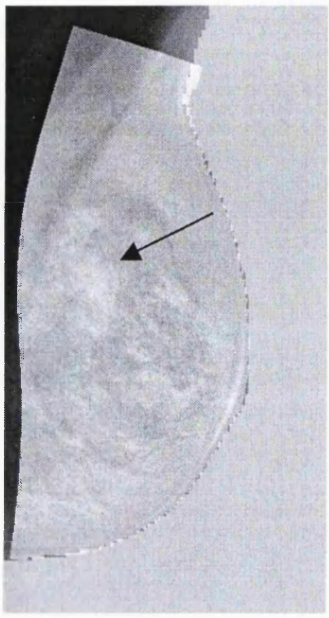


SOURCE TRANSFORMED



TRANSFORMED GRID

SUBTRACTION BEFORE



SUBTRACTION AFTER

COMMENTS

The image pair is normalised using the h_{int} representation of interesting tissue. The second mammogram is slightly less dense than the first one. Although the compression is slightly different the internal structures are significantly displaced between the two mammograms as is shown by the transformation grid. After registration, the difference image highlights a region of involution (density decrease).

List of publications

Marias, K., M., Brady, J.M , Highnam, R.P., Parbhoo, S., Seifalian, A.M., "Registration of Temporal Mammograms", *Proceedings of MIUA 99*, University of Oxford, United Kingdom, 1999.

Marias, K., Behrenbruch, C.P., Brady, M., Parbhoo, S, Seifalian, A, "Multi-scale landmark selection for improved registration of temporal mammograms", *Proceedings of the International Workshop in Digital Mammography*, Medical Physics Publishing, Toronto, Canada, June 2000.

Marias, K., Behrenbruch, C.P., Highnam, R.P., Brady, J.M., Parbhoo, S., Seifalian, A.M., "Quantifying mammographic changes in temporal HRT sequences", *Proceedings of MIUA 2000*, University College London, United Kingdom, 2000.

Marias, K., Behrenbruch, C.P., Highnam, R.P., Brady, J.M., Parbhoo, S., Seifalian, A.M., "Volume preserving elastic transformation for local breast-tissue quantification", *Proceedings of MIUA 2001*, University of Birmingham, United Kingdom, 2001.

Behrenbruch, C.P., Marias, K, Armitage, P.A., Yam, M., Moore, N., English, R.E., Brady, J.M., "Mammography-MRI 2D/3D Data Fusion for Breast Pathology Assessment", *Proceedings of MICCAI* (Medical Robotics, Imaging and Computer Assisted Intervention), Springer, Pittsburgh, USA, October 2000.

Behrenbruch, C.P., Marias, K., Armitage, P.A., Brady, J.M., Clarke, J., Moore, N., "The Generation of Simulated Mammograms from Contrast-Enhanced MRI for Surgical Planning and Postoperative Assessment", *Proceedings of the International Workshop in Digital Mammography*, Kluwer Academic Publishers, Toronto, Canada, June 2000.

Behrenbruch, C.P., Marias, K., Armitage, P.A., Yam, M., Moore, N., English, R.E., Clarke P.J., Brady, M., "Fusion of Contrast-Enhanced Breast MR and Mammographic Imaging Data", *Medical Image Analysis* (MedIA), Elsevier, 2001 (under review)

Behrenbruch, C.P., Marias, K., Armitage, P., Moore, N., Clarke, J., Brady, M., "Prone-Supine Breast MRI Registration for Surgical Visualisation", *Proceedings of MIUA 2001*, University of Birmingham, United Kingdom, 2001.

Behrenbruch, C.P., Moore, N., Marias, K., Armitage, P., Brady, M., English, R., Clarke, J., "Multimodal Data Fusion in Breast Imaging", *ECR* (European Congress of Radiology) 2001, Vienna, Austria, March 2001.

References

[1] Esteve J. Et al, "Facts and Figures of Cancer in the European Community", Technical Report, International Agency for Research on Cancer, Lyon, France, 1993.

[2] Finn J.R., "Digital Image Processing Strategies for Mammography", PhD Thesis, Centre for Biological and Medical Systems, Imperial College of Science, Technology and Medicine, University of London, 1997.

[3] Dean, P. B., " Overview of Breast Cancer Screening", *In Proceedings of Digital Mammography 96*, pp. 19-26, Elsevier, Chicago, 1998.

[4] Highnam R.P and Brady J.M. *Mammographic image processing*. Kluwer Academic Press, 1999.

[5] Gale A.G. and Cowley H.C., "Analysis of Breast Cancer Screening Results", *In Proceedings of Digital Mammography 96*, pp. 19-26, Elsevier, Chicago, 1998.

[6] Evertsz C. J.G., Bodicker A., Jan H, Hendriks C.L, Karssemeijer N., Weber U., Bohnenkamp S., Dechow D., Berger L., van Woudenberg S., Brady M., Holland R., Heinz-Otto Peitgen, "High-throughput soft-copy reading system for digital mammography in nationwide European screening mammography programs: Requirements and first solutions", *In Proceedings of RSNA*, November 2000.

[7] C.J.G. Evertsz, A. Boedicker, S. Bohnenkamp, D. Dechow, L. Berger, U. Weber, C. Beck, J.H.C.L. Hendriks, N. Karssemeijer, Michael Brady, H. Jurgens, and H.-O. Peitgen, "Soft copy reading environment for screening mammography – SCREEN", *In Proceedings of International Workshop on Digital Mammography* (IWDM), Toronto, Canada, June 2000.

[8] Andolina V.F., Lillé S.L. and Willison K.M., "Mammographic Imaging: A Practical Guide", J.B. Lippincott Company, 1992.

[9] General Electric Medical Systems, URL: http://www.gemedicalsystems.com/rad/

[10] Shapiro S., Venet W., and Strax P., "Ten- to fourteen-year effect of screening on breast cancer mortality", *J Natl Cancer Inst,* vol. 69, pp. 349-355, 1982.

[11] Kopans D.B., "Positive predictive value of mammography", *Am J Roentgenology,* vol. 158, pp. 521-526, 1992.

[12] Bird R.E., Walladce T.W., and Yankaskas B.C., "Analysis of cancers missed at the screening mammography", *Radiology,* vol. 184, pp. 613-617, 1992.

[13] Samuel K. Moore, "Better Breast Cancer Detection", *IEEE Spectrum*, Vol 38, No. 5, May 2001 .

[14] Sabel M. and Aichingert H., "Recent Developments in Breast Imaging", *Physics in Med. Biol.*, Vol 41, pp. 315-368, 1996.

[15] Webb S., "The Physics of Medical Imaging", IOP Publishing, 1995.

[16] Joseph P. Hornak, "The basics of MRI", online book, http://www.cis.rit.edu/htbooks/mri/

[17] Hayton, P., "Analysis of Contrast-Enhanced Breast MR Images", *D.Phil Thesis*,The University of Oxford,1998.

[18] Hayton P., Brady M., Tarassenko L., Moore N., "Analysis of dynamic MR breast images using a model of contrast enhancement", *Medical Image Analysis*, Oxford University Press, 207-224,1996.

[19] National Technology Alliance (organisation), Breast Cancer Initiative Home Page, http://www.nml.org/Breast Cancer.

[20] Heywang-Kobrunner S.H., Schreer I., and Dershaw D.D., In *Diagnostic breast imaging,* pp. 11-100, Thieme Stuttgart, New York, 1997.

[21] Heywang S.H., Wolf A., Pruss E., et al "MR imaging of the breast with Gd-DPTA: Use and limitations," *Radiology,* vol. 190, pp. 484-493, 1994.

[22] Wende Logan-Young, and Nancy Yanes Hoffman, "Breast Cancer: A practical Guide to Diagnosis", Vol. 1- Procedures, 1st Edition, Mt. Hope Publishing, 1994.

[23] Burcher M., "3D Freehand Breast Ultrasound: Correction of Tissue Deformation and Measures of Tumour Vascularity", Department of Engineering Science, University of Oxford, 2000.

[24] Klaus AJ, Klingensmith WC 3rd, Parker SH, Stavros AT, Sutherland JD, Aldrete KD, "Comparative value of 99mTc-sestamibi scintimammography and sonography in the diagnostic workup of breast masses", *AJR Am J Roentgenol*, 174(6):1779-83, Jun 2000.

[25] C.J. Thompson, Murphy K., Picard Y et al, "Positron Emission Mammography: A Promising Technique for Detecting Breast Cancer", IEEE Trans. On Nuclear Science, Vol 42, No. 4, pp. 1012-1017, August 1995.

[26] Buscombe J., Hill J., Parbhoo S., "Scintimammography: A guide to good practice", Gibbs Associates Limited, 1998.

[27] Dixon J.M. (editor), "ABC of Breast Diseases", BMJ Publishing Group, 1995.

[28] Orel S.G., Mendonca M.H., Reynolds C., Schnall M.D., Solin L.J., Sullivan D.C., "MR imaging of ductal carcinoma in situ", *Radiology*, 202(2):413-20, Feb 1997.

[29] Powell D.E., Sterling C.B., "Magnetic resonance imaging of the human female breast. Current status and pathologic correlations", *Pathol Annu.,* 23 Pt 1:159-94,1988

[30] Drew P.J., Chatterjee S., Turnbull L.W., Read J., Carleton P.J., Fox J.N., Monson J.R., Kerin M.J., "Dynamic contrast enhanced magnetic resonance imaging of the breast is superior to triple assessment for the pre-operative detection of multifocal breast cancer", *Ann Surg Oncol.*, 6(6):599-603, Sep 1999.

[31] Gruber G.J., Moses W.W., Derenzo S.E., Wang N.W, Beuville E., and Ho M.H., "A Discrete Scintillation Camera Module Using Silicon Photodiode Readout of CsI(Tl) Crystals for Breast Cancer Imaging", *IEEE Trans Nucl Sci,* vol. NS-45, pp. 1063-1068, 1998.

[32] Behrenbruch, C.P., Marias, K., Armitage, P.A., Brady, J.M., Clarke, J., Moore, N., "The Generation of Simulated Mammograms from Contrast-Enhanced MRI for Surgical Planning and Postoperative Assessment", *Proceedings of the International Workshop in Digital Mammography*, Toronto, Canada, Kluwer Academic Publishers, June 2000.

[33] Behrenbruch, C.P., Marias, K, Armitage, P.A., Yam, M., Moore, N., English, R.E., Brady, J.M., "Mammography-MRI 2D/3D Data Fusion for Breast Pathology Assessment", *Proceedings of MICCAI (Medical Robotics, Imaging and Computer Assisted Intervention)*, Pittsburgh, USA, Springer, October 2000.

[34] Roehrig J., Doi T., Hasegawa A. et al, "Clinical results with R2 Imagechecker system" *In Proceedings of Digital Mammography 98*, pp. 395-400, Kluwer Academic Publisher, Nijmegen, 1998.

[35] Funovics M, Schamp S, Lackner B, Wunderbaldinger P, Lechner G, Wolf G., "Computer-assisted diagnosis in mammography: the R2 ImageChecker System in detection of spiculated lesions", *Wien Med Wochenschr*, 148(14):321-4, 1998.

[36] Yam M., Highnam R.P., Brady J.M, English R., Kita Y., "Three-dimensional reconstruction of microcalcification clusters from two mammographic views", Proc International Workshop on Digital Mammography (IWDM), Toronto, Canada, 2000.

[37] Yam, M., Brady, J.M., Highnam, R.P., English, R.E. and Kita, Y., "Reconstructing microcalcification clusters in 3-D using a parameterized breast compression model", *In Proceedings of Computer Assisted Radiology and Surgery (CARS), Paris*, 1999.

[38] Rangayyan RM, Mudigonda NR, Desautels JE, "Boundary modelling and shape analysis methods for classification of mammographic masses", *Med Biol Eng Comput*, 38(5):487-96, Sep 2000.

[39] Priebe C.E., Solka J.L., Lorey R.A. et al "The application of fractal analysis to mammographic tissue classification", *Cancer Lett*, 15;77(2-3):183-9, Mar 1994.

[40] Taylor P, Fox J and Todd-Pokropek A., "A Model for Integrating Image Processing into Decision Aids for Diagnostic Radiology", *Artificial Intelligence in Medicine*; 9:205-225, 1997.

[41] Taylor P, Fox J, Todd-Pokropek A., "The development and evaluation of CADMIUM: a prototype system to assist in the interpretation of mammograms", *Medical image analysis*, 3: 321-337, 1999.

[42] Taylor P., "Decision Support for Image Interpretation: A Mammography Workstation", In Bizais Y, Barillot C and di Paola R eds., *Information Processing in Medical Imaging*, Kluwer Academic Publishers, pp. 227-239, 1995.

[43] Cerneaz N. J. and Brady J. M., "Finding curvilinear structures in mammograms", In N. Ayache, editor, *CVRMed'95*, Lecture Notes in Computer Science, pages 372 - 382, Nice, France, Springer, April 1995

[44] Antoine Maintz J.B. and Viergever M.A., "A survey of medical image registration", Medical Image Analysis, volume 2, No.1, pp 1-36, Oxford University Press, 1998.

[45] Clarkson M.J., Rueckert D., Hill D.L.G. and Hawkes D.J., "Registration of multiple video images to pre-operative data for image guided surgery", *Proceedings Medical Image Understanding and Analysis MIUA98,* pp. 73-76, 1998.

[46] Clarkson M.J., Rueckert D., Hill D.L.G. and Hawkes D.J., "Registration of multiple video images to pre-operative CT for image guided surgery", *Medical Imaging 1999,* vol. SPIE 3661, 1999.

[47] Feldmar J., Ayache N., "Rigid, Affine and Locally Affine Registration of Free-Form Surfaces", *Technical Report,* INRIA, France, 1994.

[48] Thompson, P.M., Toga, A.W., "Detection, visualization and animation of abnormal anatomic structure with a deformable probabilistic brain atlas based on random vector field transformations", *Medical Image Analysis*, Oxford University Press, 271-294, 1997.

[49] Kyriacou S.K., Davatzikos S., Zinreich J., Bryan R.N., "Nonlinear Elastic Registration of Brain Images with Tumor Pathology Using a Biomechanical Model", *IEEE Transactions on Medical Imaging*, 580-592, 1999.

[50] Colchester A.C.F., Zhao J., Holton-Tainter K.S., Henri C.J., Maitland N., Roberts P.T.E., Harris C.G., and Evans R.J., "Development and preliminary evaluation of VISLAN, a surgical planning and guidance system using intra-operative video imaging", *Medical Image Analysis*, vol. 1, 1996.

[51] Henri C.J., Colchester A.C.F., Zhao J., Hawkes D.J., Hill D.L.G., and Evans R.L., "Registration of 3-D surface data for intra-operative guidance and visualisation in frameless stereotactic neurosurgery", *In: Computer Vision, Virtual Reality and Robotics in Medicine (CVRMed '95),* ed. N. Ayache, Springer, pp. 47-69, 1995.

[52] Weese J., Penney G.P., Desmedt P., Buzug T.M., Hill D.L.G., and Hawkes D.J., "Voxel based 2D/3D registration for image guided surgery", *IEEE transactions on information technology in biomedicine,* vol. 1, pp.284-293, 1997.

[53] Grimson W.E.L., Ettinger G.J., White S.J., Lozano-Perez T., Wells III W.M., Kikinis R., "An Automatic Registration Method for Frameless Stereotaxy, Image Guided Surgery, and Enhanced Reality Visualization", *IEEE Transactions on Medical Imaging*, 129-140, 1996.

[54] Cotin, S., "Modeles anatomiques deformables en temps reel: Application a la simulation de chirurgie avec retour d'effort", *Doctoral Thesis (These)*, Universite de Nice - Sophia Antipolis, 1997.

[55] Cotin, S., Delingette, H., Ayache, N., "Real-time elastic deformations of soft tissues for surgery simulation", *Technical Report No. 3511*, INRIA, Theme 3, 1998.

[56] Edwards, P.J., Hill, D.J.G., Little, J.A., Hawkes, D.J., "A three-component deformation model for image-guided surgery", *Medical Image Analysis*, Oxford University Press, 355-367, 1998.

[57] Azar, F.S., Metaxas, D.N., Schnall, M.D., "A Finite Element Model of the Breast for Predicting Mechanical Deformation During Biopsy Procedures", *Proceedings of the IEEE Workshop on Mathematical Methods in Biomedical Image Analysis*, 314-321, 2000.

[58] Declerck J., Feldmar J., Goris M.L., Betting F., "Automatic Registration and Alignment on Template of Cardiac Stress & Rest SPECT Images", *Technical Report*, INRIA, France, 1996.

[59] Evans, A.C., Dai, W., Collins, L., Neelin, P.and Marrett, S., "Warping of a computerized 3-D atlas to match brain image volumes for quantitative neuroanatomical and functional analysis", *In Proceedings of SPIE*, Medical Imaging V: Image Processing, Leow (ed.), San Jose, CA, USA, pp. 236-246, 1991.

[60] Evans A.C., Collins D.L., Neelin P., Marrett T.S., "Correlative Analysis of Three-Dimentional Brain Images", *Registration*, 99-114, 1991.

[61] Christensen, G.E., Rabbitt, R.D. and Miller, M.I., "3D brain mapping using deformable neuroanatomy", *Physics in Medicine and Biology*, 609-618, 1994.

[62] McInerney, T., Terzopolous, D., "Deformable models in medical image analysis: a survey", *Medical Image Analysis*, Oxford University Press, 91-108, 1996.

[63] McInerney, T., Terzopoulos, D., "Deformable Models in Medical Image Analysis", *Proceedings of MMBIA '96*, IEEE, 171-180, 1996.

[64] Declerck, J., "Etude de la Dynamique Cardiaque par Analyse D'Images Tridimensionnelles", *Doctoral Thesis*, L'Universite de Nice Sophia-Antipolis, 1997.

[65] Declerck J., Feldmar J., Ayache N., "Definition of a 4D continuous polar transformation for the tracking and the analysis of LV motion", *Technical Report*, INRIA, France, 1996.

[66] Rueckert D., Hayes C., Studholme C., Summers P.E., Leach M., and Hawkes D.J., "Non-rigid registration of breast MR images using mutual information", *In 1st International Conference of Medical Imaging Computing and Computer Assisted Intervention*, pp1144-1152, Berlin: Springer-Verlag, 1998.

[67] Rueckert D., Sonoda L.I., Hayes C., Hill D.L., Leach M.O., Hawkes D.J., "Non-rigid registration using free-form deformations: application to breast MR images", *IEEE Transactions on Medical Imaging*, 712-721, 1999.

[68] Morsy, A.A., Von Ramm., O.T., "3D Ultrasound Tissue Motion Tracking Using Correlation Search", *Ultrasonic Imaging*, 151-159, 1998.

[69] Tanner C., Schnabel J.A., Chung D., Clarkson M.J., Ruecket D., Hill, D.L.G., Hawkes D.J., "Volume and shape Preservation of Enhancing Lesions when Applying Non-rigid Registration to a Time Series of Contrast Enhancing MR Breast Images", *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 1998.

[70] Woods R. P., Mazziotta J. C., and, Cherry S. R.,"MRI-PET registration with automated algorithm", *J. Computer Assisted Tomography,* vol. 16, pp.620-633, 1992.

[71] Studholme C., Hill D.L.G., and Hawkes D.J., "Automated registration of truncated MR and CT datasets of the head", In Proceedings of British Machine Vision Conference, pp. 27-36, 1995.

[72] Studholme C., Hill D.L.G., Hawkes D.J., "Automated 3-D registration of MR and CT images of the head", *Medical Image Analysis*, Oxford University Press, 163-175, 1996.

[73] Guimond, A., Roche, A., Ayache, N., Meunier, J., "Multimodal Brain Warping Using the Demons Algorithm and Adaptative Intensity Corrections", *IEEE Transactions on Medical Imaging*, 2001.

[74] Arad, N. and Reisfeld, D., "Image warping using few anchor points and radial functions", *Computer Graphics Forum*, 35-36, 1995.

[75] Little, J.A., Hill, D.L.G. and Hawkes, D.J., "Deformations Incorporating Rigid Structures", *Computer Vision and Image Understanding*, 223-232, 1997.

[76] Bookstein F., "Principal warps: Thin-plate splines and the decomposition of deformations", *IEE Transactions Pattern Analysis and Machine Intelligence*, 11, pp. 567-585, 1989.

[77] Suter D., Chen F., "Left Ventricular Motion Reconstruction Based on Elastic Vector Splines", *IEEE Transactions on Medical Imaging*, 295-305, 2000.

[78] Papademetris, X., Shi, P., Dione, D.P., Sinusas, A.J., Constable, R.T., Duncan, J.S., "Recovery of Soft Tissue Object Deformation from 3D Image Sequences using Biomechanical Models", *Technical Report 1999-01*, Image Processing and Analysis Laboratory, Yale University, 1999.

[79] Rohr, K., Stiehl, H.S., Sprengel, R., Beil, W., Buzug, T.M., Weese, J. and Kuhn, M.H., "Point-based elastic registration of medical image data using approximating thin-plate splines", *Lecture Notes in Computer Science*, Springer, 297-306, 1996.

[80] Lester H., Arridge S. R., "Summarising Fluid Registration by Thin-Plate Spline Warps with Many Landmarks", *Medical Image Analysis and Understanding (MIUA)*, 1999.

[81] Maes, F., Collingnon, A., Vandermeulen, D., Marchal, G., Suetens, P., "Multimodality image registration by maximisation of mutual information", *IEEE Transactions in Medical Imaging*, 1997.

[82] Wells W.M., Viola P., Atsumi H., Nakajima S., and Kikinis R., "Multi-modal volume registration by maximasation of mutual information", *Medical Image Analysis*, vol. 1, pp.35-51, 1996.

[83] Maintz, J.B.A., Meijering, E.H.W., Viergever, M.A., "General multimodal elastic registration based on mutual information", *In Medical Imaging: Image Processing (MI'98) (K.M. Hanson, Ed.)*, SPIE Proceedings, 1998.

.

[84] Viola P.A., "Alignment by maximisation of mutual information", PhD Thesis, MIT, 1995.

[85] Viola, P., Wells, W., "Alignment by Maximisation of Mutual Information", *International Conference in Computer Vision*, I.C.S. Press, 16-23, 1995.

[86] Gilles, S., "Description and experimentation of image matching using mutual information", *D.Phil Thesis*, The University of Oxford, 1999.

[87] G.P. Penney, J. Weese, J.A. Little, D.L.G. Hill, and D.J. Hawkes, "A comparison of similarity measures for use in 2D-3D medical image registrtation", *IEEE Transactions in Medical Imaging,* vol. 17, pp. 586-595, 1998.

[88] Roche, A., Malandain, G., Ayache, N., Prima, S., "Towards a better comprehension of similarity measues used in medical image registration", *In the Proceedings of MICCAI '99, Lecture Notes in Computer Science*, Springer-Verlag, 555-566, 1999.

[89] Roche A., Malandain G., Pennec X., and Ayache N., "The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration", in *Proceedings of Medical Imaging Computing and Computer Assisted Intervention MICCAI'98,* pp. 1115-1124, Lecture Notes in Computer Science, Oct. 1998. Electronic version: http://www.inria.fr/RRRT/RR-3378.html

[90] Horn, B.K.P. and Schunck, B.G., "Determining Optical Flow", *Artificial Intelligence*, 185-203, 1981.

[91] Thirion J.-P., "Fast Non-Rigid Matching of 3D Madical Images", *Technical Report*, INRIA, France, 1995.

[92] Thirion J.-P., "Image matching as a diffusion process: an analogy with Maxwell's demons", *Medical Image Analysis*, 243-260, 1998.

[93] Benayoun, S., "Calcul Local du Mouvement: Application a l'Imagerie Medicale Multidimensionelle", *Doctoral Thesis (These)*, U.F.R. Mathematiques de la Decision, 1994.

[94] Christensen, G.E., Miller, M.I., Vannier, M., "A 3D deformable magnetic resonance textbook based on elasticity", *Applications of Computer Vision in Medical Image Processing*, pp 153-156, 1994

[95] Bro-Nielsen M., Gramkow, C., "Fast Fluid Registration of Medical Images", *In Proceedings of VBC'96, Lecture Notes in Computer Science (Hohne and Kikinis Eds)*, Springer-Verlag, 1996.

[96] Besl, P. and McKay, N.,"A method for registration of 3-D shapes", *IEEE Trans. PAMI*, 239-256, 1992.

[97] Zhang Z., "Iterative Point Matching for Registration of Free-Form Curves and Surfaces", *the International Journal of Computer Vision*", 13(2):119-152, 1994.

[98] Feldmar, J., Declerck, J., Malandain, G., "Extension of the ICP algorithm to non-rigid intensity-based registration of 3D volumes", *Computer Vision and Image Understanding*, 1997.

[99] Michael Fitzpatrick J. and West Jay B., "Empirical Evaluation Technique in Computer Vision", ed. Kevin W. Bowyer and P. Jonathon Phillips, IEEE Computer Society Press, Los Alamitos, CA, 1998.

[100] Siew-Li Kok-Wiles, Brady M. and Hignam R., "Comparing mammogram pairs for the detection of lesions", In Proceedings of *Digital Mammography 98*, pp. 103-110, Kluwer Academic Publisher, Nijmegen 1998.

[101] Highnam R.P., Brady J.M., Shepstone B.J., "Detecting screen-film artifacts in mammography using a model-based approach", IEEE Medical Imaging, Vol. 18 No. 10, pp. 1016-1024, Oct. 1999.

[102] Siew-Li Kok-Wiles, "Comparing mammogram pairs in the detection of mammographic lesions", DPhil thesis, Robotics Research Group, Department of Engineering Science, University of Oxford, 1996.

[103] Novak R., "Transformation of the female breast during compression at mammography with special reference to the importance for localisization of a lesion", PhD thesis, Department of Diagnostic Radiology at Lakarhuset and Karolinska Sjukhuset, Sweden, 1989, Acta Radiologica Supplement 371.

[104] Vujovic N., Balic P.and Brzakovic D., "Detection of potentially cancerous signs by mammogram follow up. In K. Doi, M.L. Giger, R.M. Nishikawa, and R.A. Schmidt, editors, *Digital Mammography 96*, Chicago, Elsevier Science B.V., June 1996.

[105] Vujovic N. and Brzakovic D., "Establishing the correspondance between control points in pairs of mammographic images", *IEE Transactions on Medical Imaging*, 6, No. 10, pp. 1388-1398, October 1997.

[106] Sanchez-Ortiz G.I., Rueckert D. and Burger P., "Motion and deformation analysis of the heart using thin-plate splines and density and velocity encoded MR Images", In *16th Statistical Workshop on Image Fusion and Shape Variability*, pages 71-78, Leeds, UK, July 1996.

[107] Karssemeijer N. and Brake G.T., "Combining single view features and asymmetry for detection of mass lesions", *In Proceedings of Digital Mammography 98*, pp. 95-102, Kluwer Academic Publisher, Nijmegen, 1998.

[108] Sallam M. and Bowyer K., "Registration and difference analysis of corresponding mammogram images", *Medical Image Analysis*, 1, No. 1, pp 73-91, 1996.

[109] Asada H. and Brady M., "The curvature primal sketch", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, pp. 2-14, 1986.

[110] Coifman R.R., Meyer Y., Quake S.R. and Wickerhauser M.V., "Signal processing and compression with wavelet packets", *Proceedings of the International Conference on Wavelets*, Marseille, (Y. Meyer and S. Roques, eds.) Springer, Berlin, pp 77-93, 1992.

[111] McGregor A., "Family Doctor Guide to: The menopause and HRT", The British Medical Association, Dorling Kindersley, 2000.

[112] Hope S., Rees M. and Brockie J., "Hormone Replacement Therapy: a guide for primary care", Oxford Medical Publications, Oxford University Press 1999.

[113] Bush T.L., "The Epidemiology of Cardiovascular Disease in Postmenopausal Women", *Annals of the New York Academy of Sciences, Multidisciplinary Perspectives on Menopause*, M. Flint, F. Kronenberg, and W. Utian (eds.), vol. 592, p. 264, 1990.

[114] Komulainen M, Kroger H, Tuppurainen MT et al., "Identification of early postmenopausal women with no bone response to HRT: results of a five-year clinical trial", *Osteoporos Int*;11(3):211-8, 2000.

[115] Khastgir G., Studd J., "Hysterectomy and HRT", Martin Dunitz Ltd, 1998

[116] National Institutes of Health (NIH), National Heart, Lung, and Blood Institute (NHLBI), Women's Health Initiative (WHI), http://www.nhlbi.nih.gov/whi/index.html

[117] Schrott HG, Bittner V, Vittinghoff E, Herrington DM, Hulley S, for the HERS Research Group, "Adherence to National Cholesterol Education Program treatment goals in postmenopausal women with heart disease: the Heart and Estrogen/progestin Replacement Study (HERS): the HERS Research Group", *JAMA,* 277:1281-6, 1997.

[118] Hulley S, Grady D, Bush T, for the Heart and Estrogen/progestin Replacement Study (HERS) Research Group, et al., "Randomized trial of estrogen plus progestin for secondary prevention of coronary heart disease in postmenopausal women", *JAMA,* 280:605-13, 1998.

[119] Medical Research Council: http://www.mrc.ac.uk/welcome.htm

[120] Collaborative Group on Hormonal Factors in Breast Cancer, "Breast cancer and hormone replacement therapy: collaborative analysis of data from 51 epidemiological studies of 52705 women with breast cancer and 108411 women without breast cancer", *The Lancet*, Vol 350, No. 9084, pp. 1047-1059, October 1997.

[121] Gapstur SM, Morrow M, Sellers TA., "Hormone replacement therapy and risk of breast cancer with a favorable histology: results of the Iowa Women's Health Study", JAMA, 9;281(22):2140-1, Jun 1999.

[122] Litherland JC, Evans AJ, Wilson AR, "The effect of hormone replacement therapy on recall rate in the National Health Service Breast Screening Programme", *Clin Radiol*, 52(4):276-9, 1997.

[123] Holli K, Isola J, Cuzick J., "Low biologic aggressiveness in breast cancer in women using hormone replacement therapy", J Clin Oncol, 16(9):3115-20, Sep 1998

[124] Gajdos C, Tartter PI, Babinszki A., "Breast cancer diagnosed during hormone replacement therapy", Obstet Gynecol, 95(4): 513-8, Apr 2000.

[125] Rand T, Heytmanek G, Seifert M et al., "Mammography in women undergoing hormone replacement therapy. Possible effects revealed at routine examination", *Acta Radiol*, 38(2):228-31, 1997.

[126] Speroff L, "Postmenopausal hormone therapy and breast cancer", *Obstet Gynecol*, 87(2 Suppl):44S-54S, 1996.

[127] Lando JF, Heck KE, Brett KM., "Hormone replacement therapy and breast cancer risk in a nationally representative cohort", *Am J Prev Med*, 17(3):176-80, 1999.

[128] Beral V., Banks E., Reeves G., Wallis M., "Hormone replacement therapy and high incidence of breast cancer between mammographic screenings", *Lancet*, 349, 1103-1104, 1997.

[129] Beral V, Banks E, Reeves G, Appleby P., "Use of HRT and the subsequent risk of cancer", *J Epidemiol Biostat*, 4(3):191-210; discussion 210-5, 1999.

[130] The Million Women Study Collaborative Group, "The million women study: design and characteristics of the study population", *Breast Cancer Research*, 1, 73-80, 1999.

[131] The Million Women Study Collaborative Group, URL: http://www.icnet.uk/research/studies/mws/

[132] Kaufman Z, Garstin WI, Hayes R, Michell MJ, Baum M, "The mammographic parenchymal patterns of women on hormonal replacement therapy", *Clin Radiol*, 43(6):389-92, 1991.

[133] Sala E, Warren R, McCann J et al., "High-risk mammographic parenchymal patterns, hormone replacement therapy and other risk factors: a case-control study", *Int J Epidemiol*, 29(4):629-36, 2000.

[134] Sterns EE, Zee B, "Mammographic density changes in perimenopausal and postmenopausal women: is effect of hormone replacement therapy predictable?", *Breast Cancer Res Treat*, 59(2):125-32, 2000.

[135] Leung W, Goldberg F, Zee B, Sterns E, "Mammographic density in women on postmenopausal hormone replacement therapy", *Surgery*, 122(4):669-73; discussion 673-4, 1997.

[136] Marugg RC, van der Mooren MJ, Hendriks et al., "Mammographic changes in postmenopausal women on hormonal replacement therapy", *Eur Radiol*, 7(5):749-55, 1997.

[137] Rutter CM, Mandelson MT, Laya MB, Seger DJ, Taplin S., "Changes in breast density associated with initiation, discontinuation, and continuing use of hormone replacement therapy", *JAMA*, 10;285(2):171-6, 2001.

[138] Stomper PC, Van Voorhis BJ, Ravnikar VA, Meyer JE, "Mammographic changes associated with postmenopausal hormone replacement therapy: a longitudinal study", *Radiology*, 174(2):487-90, 1990.

[139] McNicholas MM, Heneghan JP, Milner MH et al., "Pain and increased mammographic density in women receiving hormone replacement therapy: a prospective study", *AJR Am J Roentgenol*, *163*(2):311-5, 1994.

[140] Cyrlak D, Wong CH, "Mammographic changes in postmenopausal women undergoing hormonal replacement therapy", *AJR Am J Roentgenol*, 161(6):1177-83, 1993.

[141] Del Favero C, Rossini G, Tufarulo L et al., "Mammography changes associated with hormone replacement therapy in post-menopausal patients", *Radiol Med (Torino)* 93(3):210-3, 1997.

[142] Reichenbach JR, Przetak C, Klinger G, Kaiser WA, "Assessment of breast tissue changes on hormonal replacement therapy using MRI: a pilot study, *J Comput Assist Tomogr*, 23(3):407-13, 1999.

[143] Porfiri LM, Costanza L, De Felice C et al., "A mammographic evaluation of the morphostructural variations of the breast during hormone-replacement therapy in the postmenopause", *Radiol Med (Torino)*, 95(6):573-6, 1998.

[144] Lundstrom E, Wilczek B, von Palffy Z et al., "Mammographic breast density during hormone replacement therapy: differences according to treatment", *Am J Obstet Gynecol*, 181(2):348-52, 1999.

[145] Ozdemir A, Konus O, Nas T et al., "Mammographic and ultrasonographic study of changes in the breast related to HRT", *Int J Gynaecol Obstet*, 67(1):23-32, 1999.

[146] Kavanagh AM, Mitchell H, Giles GG, "Hormone replacement therapy and accuracy of mammographic screening", *Lancet*, 22;355(9200):270-4, 2000.

[147] Berkowitz JE, Gatewood OM, Goldblum LE, Gayler BW, "Hormonal replacement therapy: mammographic manifestations", *Radiology*, 174(1):199-201, 1990.

[148] Harvey JA, "Use and cost of breast imaging for postmenopausal women undergoing hormone replacement therapy", *AJR Am J Roentgenol*, 172(6):1615-9, 1999.

[149] Parbhoo S.P., Seifalian A.M, "Distribuition of breast skin blood flow in patients with breast cancer", *The Breast*, 7:201-205, 1998.

[150] Maculotti L, Gandini F, Pradella P, Andrea Ponti, "Bilateral breast carcinoma. 12 years' experience", *Minerva Chir*, 51(1-2):33-7, 1996.

[151] Azzena A, Zen T, Ferrara A, Brunetti V, Vasile C, Marchetti M, "Risk factors for breast cancer. Case-control study results", *Eur J Gynaecol Oncol*, 15(5):386-92, 1994.

[152] Litherland JC, Stallard S, Hole D, Cordiner C, "The effect of hormone replacement therapy on the sensitivity of screening mammograms", *Clin Radiol*, 54(5):285-8, 1999.

[153] Byng JW, Yaffe MJ, Jong RA, Shumak RS, Lockwood GA, Tritchler DL, Boyd NF, "Analysis of mammographic density and breast cancer risk from digitized mammograms", *Radiographics*, 18(6):1587-98, 1998.

[154] Byng, J.W., Yaffe, M.J, .Lockwood, G., Little, L., Tritchler D., Boyd, N.F. "Automated analysis of mammographic densities and breast carcinoma risk: Results from the Canadian National Breast Screening Study", *Cancer*, 80:(1) 1997.

[155] Byng, J.W., Boyd, N.F., Fishell, E.K., Jong, R.A. and Yaffe, M.J., "Automated Analysis of Mammographic Densities", *Physics in Medicine and Biology*, 41 909-923, 1996.

[156] van Gils CH, Hendriks JH, Otten JD, Holland R, Verbeek AL, "Parity and mammographic breast density in relation to breast cancer risk: indication of interaction", *Eur J Cancer Prev*, 9(2):105-11, 2000.

[157] Persson I, Thurfjell E, Holmberg L., "Effect of estrogen and estrogen-progestin replacement regimens on mammographic breast parenchymal density", *J Clin Oncol*, 15(10):3201-7, Oct 1997.

[158] Laya MB, Gallagher JC, Schreiman JS et al., "Effect of postmenopausal hormonal replacement therapy on mammographic density and parenchymal pattern", *Radiology*, 196(2):433-7, 1995.

[159] Oza AM, Boyd NF, "Mammographic parenchymal patterns: a marker of breast cancer risk", *Epidemiol Rev*, 15(1):196-208, 1993.

[160] Behrenbruch, C.P., "Image fusion for the management of breast disease", *D.Phil Thesis*, The University of Oxford, 2001.

[161] Behrenbruch, C.P., Moore, N., Marias, K., Armitage, P., Brady, M., English, R., Clarke, J., "Multimodal Data Fusion in Breast Imaging", *ECR (European Congress of Radiology)* 2001, Vienna, Austria, March 2001.

[162] Kita, Y., Highnam, R. and Brady, M., "Correspondence between different view breast X-rays using a simulation of breast deformation", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 700-707,1998.

[163] Jameson J.L. (editor), "Principles of Molecular Medicine", Humana Press Inc., New Jersey USA, 1998.

[164] F.H. Martini, "Fundamentals of Anatomy and Physiology", 3rd edition, Prentice-Hall International, 1995.

[165] Yam M., Highnam R.P., Brady J.M, English R., "Detecting calcifications using the h_{int} representation", in proceedings *Proc. Computer Assisted Radiology and Surgery* (CARS), 373-377. 1999.

[166] Poissonnier M. and Brady J.M., " Noise equalisation for microcalsification detection", *Proc International Workshop on Digital Mammography* (IWDM), Toronto, Canada, 2000.

[167] D'Astous, F. and Jernigan, M.E., "Texture discrimination based on detailed measures of the power spectrum", *In proceedings of the 7th International Conference on Pattern Recognition*, pp83-86, 1984.

[168] Haralick, R.M., Shanmugam, K. and Dinstein, I., "Texture features for image classification", *IEEE SMC*, 3(6), pp610-621, 1973.

[169] Turner, M.R., "Texture discrimination by Gabor functions", *Biological Cybernetics*, 55, pp71-82, 1986.

[170] Bovik, A.C., Clark, M. and Giesler, W.S., "Multichannel texture analysis using localized spatial filters", *IEEE PAMI*, 12, pp. 55-73, 1990.

[171] Koenderink, J.J., "The structure of images", *Biological Cybernetics*, 50, pp363-370, 1984.

[172] Witkin, A.P. ,"Scale-space filtering", *In 8th International Joint Conference on Artificial Intelligence*, Karlsruhe, Germany, 1983.

[173] Burt, P.J. and Adelson, E.H., "The Laplacian pyramid as a compact image code", *IEEE Transactions on Communication*, 31(4), pp552-540, 1983.

[174] Daubechies, I., Grossman, A. and Meyer, Y. ,"Painless nonorthogonal expansions", *Journal of Mathematical Physics*, 27 , pp1271-1283, 1986.

[175] Stein, E., "Singular Integrals and the Differentiability Property of Functions", *Princeton University Press*, Princeton NJ, 1979.

[176] Marr, D., "Vision: A Computational Investigation Into the Human Representation and Processing of Visual Information", *W.H. Freeman*, San Francisco, 1982.

[177] Grossman, A. and Morlet, J., "Decomposition of functions into wavelets of constant shape and related transforms", *In Mathematics and Physics, Lectures on Recent Results*, L. Streit (ed.), World Scientific Publishing (Singapore), 1985.

[178] Mallat ,S.G., "A theory for multiresolution signal decomposition: The wavelet decomposition", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11, pp674-693, 1989.

[179] Mallat, S.G., "Multiresolution approximation and wavelet orthonormal basis of L^2(R)", *Transactions of the AMS*, 315, pp69-87, 1989.

[180] Daubechies, I., "Ten Lectures on Wavelets", *CBMS-NSF Regional Conference Series in Applied Mathematics (volume 61)*. SIAM Press, Philadelphia, 1992.

[181] Chambolle, A., DeVore, R.A., Lee, N-Y. and Lucier, B.J., "Nonlinear Wavelet Image Processing: Variational Problems, Compression and Noise Removal through Wavelet Shrinkage", *IEEE Transactions on Image Processing*, 5(7), pp319-335, 1998.

[182] Coifman, R., R. and Wickerhauser, M.V., "Wavelets and adapted waveform analysis", *In Wavelets: Mathematics and Applications*, Benedetto & Frazier (ed.), Studies in Advanced Mathematics, CRC Press, Boca Raton, Florida, 1992.

[183] Coifman, R.R., Meyer, Y. and Wickerhauser, M.V., "Adapted waveform analysis, wavelet-packets and applications", *In Proceedings of ICIAM '91*, pp 41-50, SIAM Press, Philadelphia, 1991.

[184] Mallat, S., "A Wavelet Tour of Signal Processing",. *Academic Press*, San Diego, 1998.

[185] Wickerhauser, M.V., "Adapted Wavelet Analysis from Theory to Software", *IEEE Press*, AK Peters, 1994.

[186] Coifman, R.R, and Wickerhauserm, M.V., "Entropy based algorithms for best basis selection", *IEEE Transactions on Information Theory*, 32, pp 712-771, 1992.

[187] Pincus, S.M., "Approximate Entropy: statistical properties and applications" *Commun. Statist.-Theory Meth*, 21, pp 3061-3077, 1992.

[188] Pincus, S.M., "Not all (possibly) "random" sequences are created equal", *Proc. Natl. Acad. Sci.*, 94, pp3513-3518, 1997.

[189] Sweldens, W., "The lifting scheme: A custom-design construction of biorthogonal wavelets", *Appl. Comput. Harmon. Anal.*, 3(2), pp186-200, 1996.

[190] Holschneider, M., "Wavelets: An Analysis Tool", *Oxford Mathematical Monographs*, Oxford University Press, 1995.

[191] Zhu, S.C. and Yuille, A., "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation", *IEEE Trans. PAMI*, 18(9), pp884-900, 1996.

[192] Ebell M., H., Web-based course to Information Mastery, Michigan State University, http://poems.msu.edu/InfoMastery/, 1999.