Registration and Lesion Classification of Contrast-Enhanced Magnetic Resonance Breast Images

Christine Tanner

A dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
of the
University of London.

October 2005

Computational Imaging Science Group,
Division of Imaging Sciences,
Guy’s, King’s and St. Thomas’ School of Medicine,
King’s College London
Abstract

Dynamic contrast-enhanced (DCE) magnetic resonance (MR) mammography provides information about tissue vascularity and permeability, which cannot be obtained by X-ray mammography. Initial DCE MR screening studies of young, high-risk women have reported a greatly improved sensitivity at a slightly reduced specificity compared to X-ray mammography. A frequent, but often ignored error source is patient motion. Many registration algorithms for DCE MR mammography have previously been developed, but their quantitative evaluation has been almost non-existent.

The first objective of this thesis is to develop a technique for the validation of registration algorithms for correcting for patient motion in DCE MR mammography. The second objective is to improve the classification of breast lesions by creating a computer aided diagnosis (CAD) system for DCE MR mammography which includes registration.

This novel validation method is based on simulating physically plausible deformations by biomechanical computer models using finite element methods. Firstly, the influence of various parameters on the model’s accuracy was assessed using 2 volunteers. The model’s accuracy was more sensitive to both the change in boundary condition and Poisson’s ratio rather than the elastic properties. Suitable configurations improved the average accuracy from 6.6mm to 2.1mm. Secondly, plausible breast deformations were simulated for 10 patients and the performance of several registration configurations was then optimized for 5 of these patients. The better configurations used normalized mutual information, volume preservation and relatively coarse control point spacing. These reduced the mean registration error of the remaining 5 patients from 1.40mm to 0.45mm. Finally, a CAD system which included image registration and a new segmentation refinement method was developed. This achieved an area under the receiver operator characteristics curve of 0.86 for leave-one-out tests. The classification performance for rigidly registered images was statistically significantly better than for the original images.

This thesis showed for the first time that including volume-preserving non-rigid registration into a CAD system for DCE MR mammography improves classification; that volume preservation increases the accuracy of flexible DCE MR mammography registrations; and that elastic properties were less important than boundary conditions and Poisson’s ratios.
Acknowledgements

As a girl, I loved jigsaws. However having realized, that rebuilding what others have taken apart is pointless, I embarked on the second best thing - to use science to explore a puzzle of nature. Many have helped me with this task and made this thesis possible.

I am most grateful to my first supervisor David Hawkes. His advice, inspiration, encouragement and optimism helped me through the roller coaster ride of research. Many thanks to my second supervisor, Derek Hill, for his constructive criticism and for the joy of experiencing his sharp wit. The people in the Computer Imaging Science Group have been extremely supportive. Sincere thanks to all of you, and especially to Matthew Clarkson and Andy Castellano Smith for getting me kick-started; to Julia Schnabel for her help with software development, experimental design and proof-reading; to David Atkinson, John Hipwell and Kate McLeish, for sharing a quiet office, for scanning volunteers and for answering my many questions; to Philipp Bachelor for mathematical advice; to Bill Crum for his absorbing discussions; to Ron Gaston for keeping the bureaucracy at bay; and to the condor cluster of more than 30 Linux machines for tirelessly working on my experiments.

I would like to thank Martin Leach, Michael Khazen, Preminda Kessar, Andreas Degenhardt, Mark White and Carmel Hayes from the Institute of Cancer Research for their support in this collaborative project. I greatly benefited from the expertise of clinicians from Guy’s and St. Thomas’ Hospital, and in particular thanks to Corrado D’Arrigo, Nicholas Beechey-Newman, Sarah McWilliams, Evelyn Sanderson and Annette Jones. The invaluable advice given by Rodney Hose, Stephen Keevil and Rebecca Land was much appreciated. I am much obliged to the volunteers and patients who provided me with experimental data. Without them, nothing could have been achieved.

I have to say thank-you to all my friends and family, especially to the runners and dancers for all the stress reduction; to my friend Elaine for making me laugh and for proof-reading; and to my parents and sisters for all their encouragement over the years. Last but not least, I am greatly indebted to my husband Rudy, for his unwavering support during the many long days which went into this endeavour and for his extensive library.
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2D</td>
<td>2-dimensional</td>
</tr>
<tr>
<td>3D</td>
<td>3-dimensional</td>
</tr>
<tr>
<td>4D</td>
<td>4-dimensional</td>
</tr>
<tr>
<td>AIR</td>
<td>automatic image registration</td>
</tr>
<tr>
<td>AUC</td>
<td>area under the ROC curve</td>
</tr>
<tr>
<td>BC</td>
<td>boundary condition</td>
</tr>
<tr>
<td>BE</td>
<td>bending energy</td>
</tr>
<tr>
<td>CAD</td>
<td>computer aided diagnosis</td>
</tr>
<tr>
<td>CC</td>
<td>correlation coefficient</td>
</tr>
<tr>
<td>CE</td>
<td>contrast-enhanced</td>
</tr>
<tr>
<td>CR</td>
<td>correlation ratio</td>
</tr>
<tr>
<td>CRE</td>
<td>consistency registration error</td>
</tr>
<tr>
<td>CT</td>
<td>computed tomography</td>
</tr>
<tr>
<td>DCE</td>
<td>dynamic contrast-enhanced</td>
</tr>
<tr>
<td>DCIS</td>
<td>ductal carcinoma in situ</td>
</tr>
<tr>
<td>DCT</td>
<td>discrete cosine transform</td>
</tr>
<tr>
<td>EBS</td>
<td>elastic body spline</td>
</tr>
<tr>
<td>FEM</td>
<td>finite element method</td>
</tr>
<tr>
<td>FFD</td>
<td>free-form deformation</td>
</tr>
<tr>
<td>FOV</td>
<td>field of view</td>
</tr>
<tr>
<td>Gd-DTPA</td>
<td>gadolinium diethylenetriaminepentaacetic acid</td>
</tr>
<tr>
<td>GM</td>
<td>geometric model</td>
</tr>
<tr>
<td>HRT</td>
<td>hormone replacement therapy</td>
</tr>
<tr>
<td>JE</td>
<td>joint entropy</td>
</tr>
<tr>
<td>LCIS</td>
<td>lobular carcinoma in situ</td>
</tr>
<tr>
<td>MAP</td>
<td>maximum a posteriori probability</td>
</tr>
<tr>
<td>MARIBS</td>
<td>magnetic resonance imaging for breast screening</td>
</tr>
<tr>
<td>MI</td>
<td>mutual information</td>
</tr>
<tr>
<td>ML</td>
<td>maximum likelihood</td>
</tr>
<tr>
<td>ME</td>
<td>membrane energy</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>MM</td>
<td>material model</td>
</tr>
<tr>
<td>MR(I)</td>
<td>magnetic resonance (imaging)</td>
</tr>
<tr>
<td>NHS</td>
<td>national health system</td>
</tr>
<tr>
<td>NMI</td>
<td>normalised mutual information</td>
</tr>
<tr>
<td>PCG</td>
<td>preconditioned conjugate gradient</td>
</tr>
<tr>
<td>PDE</td>
<td>partial differential equations</td>
</tr>
<tr>
<td>PIU</td>
<td>partitioned intensity uniformity</td>
</tr>
<tr>
<td>RIU</td>
<td>ratio image uniformity</td>
</tr>
<tr>
<td>RF</td>
<td>radiofrequency</td>
</tr>
<tr>
<td>ROC</td>
<td>receiver operating characteristics</td>
</tr>
<tr>
<td>ROI</td>
<td>region of interest</td>
</tr>
<tr>
<td>SNR</td>
<td>signal to noise ratio</td>
</tr>
<tr>
<td>SPM</td>
<td>spatial parametric mapping</td>
</tr>
<tr>
<td>SSD</td>
<td>sum of squared differences</td>
</tr>
<tr>
<td>T1</td>
<td>longitudinal relaxation</td>
</tr>
<tr>
<td>T2</td>
<td>transverse relaxation</td>
</tr>
<tr>
<td>TE</td>
<td>echo time</td>
</tr>
<tr>
<td>TPS</td>
<td>thin plate spline</td>
</tr>
<tr>
<td>TR</td>
<td>repetition time</td>
</tr>
<tr>
<td>TRE</td>
<td>target registration error</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>VIR</td>
<td>variance of intensity ratio</td>
</tr>
<tr>
<td>VP</td>
<td>volume preservation</td>
</tr>
</tbody>
</table>
List of Definitions

**Angiogenesis** is the formation of new blood vessels [Martin, 1998].

**Axial** is also called transverse. The axial plane is situated at right angle to the long axis of the body or an organ [Martin, 1998].

**Benign** tumours do not invade and destroy the tissue in which they originate or spread to distant sides in the body [Martin, 1998].

**Biopsy** is the removal of a small piece of living tissue from the body for microscopic examination [Martin, 1998].

**Coronal** plane divides the body into front and back parts.

**Cyst** is a closed cavity filled with fluid.

**Deformation** is an alteration in the shape or dimensions of an object as a result of the application of stress to it (Wikipedia).

**Ducts** are tubelike structures or channels, especially for carrying glandular secretions [Martin, 1998].

**Ductal carcinoma in situ (DCIS)** consists of ductual cancer cells that have not grown outside of their site of origin.

**Enhancement** refers, in the context of CE MR mammography, to the increase in MR signal due to the administration of a paramagnetic agent.

**Fascia** is a sheet or band of fibrous connective tissue separating or binding together muscles and organs [Martin, 1998].

**Gland** is an organ or a group of cells that is specialized for synthesizing and secreting certain fluids [Martin, 1998].

**Gold standard** is defined as the best example of its kind.

**Ground truth** is defined, in the analysis of aerial photographs, as the facts that are found when people actually visit the location. For image registration, it is the true correspondence between images.
**Interstitial space** is a small space in a tissue or between parts of the body, which is in the breast equivalent to the extravascular and extracellular space.

**Invasive** (or infiltrating) cancers have spread into neighbouring tissue, i.e. have grown through the basement membrane.

**Kinetics** is the study of the rates at which chemical reactions and biological processes proceed [Walker, 1995].

**Lactation** is the secretion of milk by the mammary glands of the breasts [Martin, 1998].

**Lobe** is a major division of an organ or part of an organ, especially one having a rounded form and often separated from other lobes by fissures or bands of connective tissue [Martin, 1998].

**Lobule** is a subdivision of a part or organ that can be distinguished from the whole by boundaries that are visible with or without a microscope [Martin, 1998].

**Malignant** tumours invade and destroy the tissue in which they originate and can spread to other sides in the body via the bloodstream and the lymphatic system [Martin, 1998].

**Mamma** means breast.

**Mammary gland** is the milk-producing gland of female mammals [Martin, 1998].

**Morphology** is the study of the structure and forms of organisms, as opposed to the study of their function [Walker, 1995].

**Necrosis** means dead tissue.

**Pixel** is a picture element, representing the smallest discrete component of a 2D image.

**Prone** means lying on the front.

**Registration** is the process of determining the spatial correspondence between features in images or real world objects.

**Sagittal** plane divides the body into left and right side.

**Sensitivity** measures the probability of detecting the disease when the disease exists.

**Specificity** measures the probability of identifying that there is no disease when the disease is absent.
**Spline** is a piecewise polynomial function that can have a locally very simple form, yet at the same time be globally flexible and smooth. [Weisstein, 2003]

**Symptom** is an indication of a disease noticed by the patient himself [Martin, 1998].

**Tumour** is an abnormal mass of tissue.

**Voxel** is a volume element, representing the smallest discrete component of a 3D image.
Notation

Notation for 3D Registration

- \( \mathbf{x} \) a point with 3D coordinates \([x_1 \ x_2 \ x_3]^T\), where \(\mathbf{x}^T\) denotes the transpose of \(\mathbf{x}\)
- \( A \) reference image with domain \(\{\mathbf{x} \mid 0 \leq x_d < X_d\}\) for \(d = 1, 2, 3\)
- \( B \) source image which should be aligned with \( A \)
- \( \{\mathbf{x}^{(p)}\} \) a set of points with 3D coordinates \([x_{1p} \ x_{2p} \ x_{3p}]^T\)
- \( A(\mathbf{x}) \) intensity value of image \( A \) at point \( \mathbf{x} \)
- \( \mathbf{T} \) geometric transformation, mapping point \( \mathbf{x} \) in \( A \) to point \( \mathbf{T}(\mathbf{x}) \) in \( B \)
- \( T_d(\mathbf{x}) \) \(d\)th coordinate of point \( \mathbf{T}(\mathbf{x}) \), i.e. \( \mathbf{T}(\mathbf{x}) = [T_1(\mathbf{x}) \ T_2(\mathbf{x}) \ T_3(\mathbf{x})]^T \)
- \( \mathbf{u}(\mathbf{x}) \) displacement vector at point \( \mathbf{x} \): \( \mathbf{u}(\mathbf{x}) = \mathbf{T}(\mathbf{x}) - \mathbf{x} \)
- \( u_d(\mathbf{x}) \) \(d\)th component of displacement vector \( \mathbf{u}(\mathbf{x}) \): \( u_d(\mathbf{x}) = T_d(\mathbf{x}) - x_d \)
- \( |\mathbf{u}| \) length of vector \( \mathbf{u} \): \( |\mathbf{u}| = \sqrt{\sum_d u_d^2} \)
- \( B^T \) image \( B \) transformed by \( \mathbf{T} \)
- \( \mathcal{A} \) discrete domain of image \( A \), sampled at \( \delta_d = \frac{X_d}{I_d} \)
- \( \mathbf{i} \) voxel index of \( \mathbf{x} \in A \) with \( \mathbf{i} = [i_1 \ i_2 \ i_3]^T \), \( i_d = \lfloor \frac{x_d}{\delta_d} \rfloor \), \( i_d \in \{0, 1, ..., I_d - 1\} \)
- \( \Omega^T \) discrete overlap domain of image \( A \) and \( B^T \), i.e. \( \mathcal{A} \cap B^T \)
- \( \mathcal{N}^T \) number of voxels of image overlap \( \Omega^T \)
- \( \overline{\mathcal{A}} \) mean value of \( A \) for overlap domain: \( \overline{\mathcal{A}} = \frac{\sum_{\Omega^T} A(\mathbf{x})}{\mathcal{N}^T} \)
- \( Var(A) \) variance of \( A \) for overlap domain: \( Var(A) = \frac{\sum_{\Omega^T} (A(\mathbf{x}) - \overline{\mathcal{A}})^2}{\mathcal{N}^T} \)
- \( Cov(A, B^T) \) covariance of \( A \) and \( B^T \) for overlap domain:
  \[
  Cov(A, B^T) = \frac{\sum_{\Omega^T} (A(\mathbf{x}) - \overline{\mathcal{A}})(B(\mathbf{T}(\mathbf{x})) - \overline{B^T})}{\mathcal{N}^T}
  \]
Notation for Continuum Mechanics

The common notation for non-linear continuum mechanics is used in this thesis [Holzapfel, 2001]. This includes that uppercase letters will denote entities of the reference configuration \( \Omega_0 \) while lowercase letters refer to entities of the current configuration \( \Omega \). The terms are defined on the referenced page.

The concept of tensors is important in the field of mechanics. Tensors are a generalization of vectors and matrices. Tensors are multidimensional arrays which are specified with respect to a given coordinate system and are able to undergo transformation to other coordinate systems. A 3D \( n \)th order tensor is defined by \( 3^n \) components and can be described by \( n \) indices.

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathcal{B} )</td>
<td>continuum body</td>
<td>92</td>
</tr>
<tr>
<td>( \Omega_0 )</td>
<td>geometrical space occupied by ( \mathcal{B} ) at time 0 (reference configuration)</td>
<td>92</td>
</tr>
<tr>
<td>( \Omega )</td>
<td>geometrical space occupied by ( \mathcal{B} ) at time ( t ) (current configuration)</td>
<td>92</td>
</tr>
<tr>
<td>( \mathbf{X} )</td>
<td>position in ( \Omega_0 )</td>
<td>92</td>
</tr>
<tr>
<td>( \mathbf{x} )</td>
<td>position in ( \Omega )</td>
<td>92</td>
</tr>
<tr>
<td>( \Phi(\mathbf{X},t) )</td>
<td>motion of body ( \mathcal{B} ), mapping position ( \mathbf{X} ) in ( \Omega_0 ) to position ( \mathbf{x} ) in ( \Omega )</td>
<td>92</td>
</tr>
<tr>
<td>( \mathbf{U}(\mathbf{X},t) )</td>
<td>displacement at position ( \mathbf{X} ): ( \mathbf{U}(\mathbf{X},t) = \mathbf{x}(\mathbf{X},t) - \mathbf{X} )</td>
<td>93</td>
</tr>
<tr>
<td>( \tilde{\mathbf{U}} )</td>
<td>prescribed displacement (boundary condition) for ( \Omega_0 )</td>
<td>97</td>
</tr>
<tr>
<td>( \tilde{\mathbf{u}} )</td>
<td>displacement at node position</td>
<td>106</td>
</tr>
<tr>
<td>( \mathbf{F}(\mathbf{X},t) )</td>
<td>deformation gradient tensor</td>
<td>93</td>
</tr>
<tr>
<td>( \mathbf{R} )</td>
<td>rotation tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{U} )</td>
<td>right stretch tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{v} )</td>
<td>left stretch tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{C} )</td>
<td>right Cauchy-Green deformation tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{d} )</td>
<td>left Cauchy-Green deformation tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \lambda_i )</td>
<td>( i )th principle stretch: ( i )th eigenvalue of ( \mathbf{U} ) or ( \mathbf{v} )</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{E} )</td>
<td>Green-Lagrange strain tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{e} )</td>
<td>Euler-Almansi strain tensor</td>
<td>94</td>
</tr>
<tr>
<td>( \mathbf{f} )</td>
<td>surface force</td>
<td>95</td>
</tr>
<tr>
<td>( \mathcal{B} )</td>
<td>body force of reference configuration</td>
<td>97</td>
</tr>
<tr>
<td>( \mathbf{b} )</td>
<td>body force of current configuration</td>
<td>95</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>Cauchy stress tensor</td>
<td>95</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>------</td>
</tr>
<tr>
<td>$\mathbf{T}$</td>
<td>traction vector of reference configuration</td>
<td>95</td>
</tr>
<tr>
<td>$\mathbf{t}$</td>
<td>traction vector of current configuration</td>
<td>95</td>
</tr>
<tr>
<td>$\mathbf{T}$</td>
<td>prescribed traction (boundary condition) for $\Omega_0$</td>
<td>97</td>
</tr>
<tr>
<td>$\mathbf{N}$</td>
<td>unit normal vector of reference configuration</td>
<td>96</td>
</tr>
<tr>
<td>$\mathbf{n}$</td>
<td>unit normal vector of current configuration</td>
<td>95</td>
</tr>
<tr>
<td>$\mathbf{P}$</td>
<td>first Piola-Kirchhoff stress tensor</td>
<td>95</td>
</tr>
<tr>
<td>$\mathbf{S}$</td>
<td>second Piola-Kirchhoff stress tensor</td>
<td>96</td>
</tr>
<tr>
<td>$J$</td>
<td>volume ratio: $J=\det(\mathbf{F}(\mathbf{X}, t))$</td>
<td>96</td>
</tr>
<tr>
<td>$\delta W_{\text{int}}$</td>
<td>internal virtual work</td>
<td>98</td>
</tr>
<tr>
<td>$\delta W_{\text{ext}}$</td>
<td>external virtual work</td>
<td>98</td>
</tr>
<tr>
<td>$\mathbf{Z}$</td>
<td>elastic tensor of reference configuration</td>
<td>100</td>
</tr>
<tr>
<td>$\mathbf{z}$</td>
<td>elastic tensor of current configuration</td>
<td>100</td>
</tr>
<tr>
<td>$E$</td>
<td>Young’s modulus</td>
<td>101</td>
</tr>
<tr>
<td>$v$</td>
<td>Poisson’s ratio</td>
<td>101</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Shear modulus</td>
<td>101</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Bulk modulus</td>
<td>113</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>strain energy function</td>
<td>101</td>
</tr>
<tr>
<td>$I_a(\mathbf{C})$</td>
<td>$a$th invariant of $\mathbf{C}$</td>
<td>102</td>
</tr>
<tr>
<td>$\mathbf{H}$</td>
<td>element shape function</td>
<td>106</td>
</tr>
<tr>
<td>$\mathbf{L}$</td>
<td>natural coordinates of element</td>
<td>106</td>
</tr>
<tr>
<td>$\mathbf{K}$</td>
<td>stiffness matrix</td>
<td>110</td>
</tr>
<tr>
<td>$M$</td>
<td>number of elements</td>
<td>108</td>
</tr>
<tr>
<td>$N$</td>
<td>number of nodes</td>
<td>108</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
<td>Description</td>
</tr>
<tr>
<td>------------</td>
<td>-----------------------------------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>( \mathbf{u} \cdot \mathbf{v} )</td>
<td>( \sum_a u_a v_a )</td>
<td>contraction (or dot product) of two vectors</td>
</tr>
<tr>
<td>( \mathbf{U} : \mathbf{V} )</td>
<td>( \sum_a \sum_b U_{ab} V_{ab} )</td>
<td>double contraction of two second-order tensors</td>
</tr>
<tr>
<td>( \text{sym}(\mathbf{V}) )</td>
<td>( \frac{1}{2}(\mathbf{V} + \mathbf{V}^T) )</td>
<td>symmetric tensor of ( \mathbf{V} )</td>
</tr>
<tr>
<td>( \text{tr}(\mathbf{V}) )</td>
<td>( \sum_a V_{aa} )</td>
<td>trace of tensor ( \mathbf{V} )</td>
</tr>
<tr>
<td>( \text{Div} \mathbf{V} )</td>
<td>( \sum_B \frac{\partial V_{ab}}{\partial X_B} )</td>
<td>divergence of tensor ( \mathbf{V} ) in reference configuration</td>
</tr>
<tr>
<td>( \text{div} \mathbf{V} )</td>
<td>( \sum_b \frac{\partial V_{ab}}{\partial x_b} )</td>
<td>divergence of tensor ( \mathbf{V} ) in current configuration</td>
</tr>
<tr>
<td>( (\text{Grad} \mathbf{U})_{aB} )</td>
<td>( \frac{\partial U_a}{\partial X_B} )</td>
<td>displacement gradient in reference configuration</td>
</tr>
<tr>
<td>( (\text{grad} \mathbf{u})_{ab} )</td>
<td>( \frac{\partial u_a}{\partial x_b} )</td>
<td>displacement gradient in current configuration</td>
</tr>
</tbody>
</table>

Chapter 3 uses the common notation for non-linear continuum mechanics. Some of these terms can be related to the common notation for image registration that was used in Chapter 2. Assuming the reference image \( \mathbf{A} \) captured the reference configuration \( \Omega_0 \) and the source image \( \mathbf{B} \) captured the current configuration \( \Omega \), then the following terms are equivalent:

<table>
<thead>
<tr>
<th>Chap. 2</th>
<th>Chap. 3</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbf{x} )</td>
<td>( \mathbf{X} )</td>
<td>position in reference (( \mathbf{A} ) resp. ( \Omega_0 ))</td>
</tr>
<tr>
<td>( \mathbf{T}(\mathbf{x}) )</td>
<td>( \mathbf{x} )</td>
<td>corresponding position in source/current (( \mathbf{B} ) resp. ( \Omega ))</td>
</tr>
<tr>
<td>( \mathbf{T} )</td>
<td>( \Phi )</td>
<td>transformation resp. motion</td>
</tr>
<tr>
<td>( \mathbf{u}(\mathbf{x}) )</td>
<td>( \mathbf{U}(\mathbf{X}, t) )</td>
<td>displacement with respect to reference (( \mathbf{A} ) resp. ( \Omega_0 ))</td>
</tr>
<tr>
<td>( \nabla \mathbf{T}(\mathbf{x}) )</td>
<td>( \mathbf{F}(\mathbf{X}, t) )</td>
<td>deformation gradient</td>
</tr>
</tbody>
</table>
Contents

1 Introduction

1.1 Breast Cancer Detection .......................................... 23
1.2 Breast Anatomy .................................................. 25
1.3 Breast Cancer .................................................... 27
1.4 DCE MR Mammography .......................................... 28
  1.4.1 MRI Basic Principles ........................................ 28
  1.4.2 Contrast-Enhanced Image Sequence .......................... 30
  1.4.3 Interpretation of DCE MR Mammography .................... 31
  1.4.4 Prospective DCE MR Mammography Studies ................ 33
1.5 Patient Motion and Registration ................................ 36
1.6 Computer Aided Diagnosis System ............................... 39
1.7 Thesis Structure .................................................. 40
1.8 Contributions of this Research .................................. 41
1.9 Software ......................................................... 42

2 Review of DCE MR Mammography Registration ................. 43

2.1 Introduction ....................................................... 43
2.2 Transformation Models ........................................... 47
  2.2.1 Global Models .............................................. 48
  2.2.2 Semi-Local Models ......................................... 56
  2.2.3 Local Models ............................................... 57
  2.2.4 Summary ................................................... 64
2.3 Image Matching Criteria .......................................... 65
  2.3.1 Feature Based Similarity Measures .......................... 65
  2.3.2 Voxel Based Similarity Measures ............................ 67
  2.3.3 Summary ................................................... 71
2.4 Optimization Strategies ......................................... 71
  2.4.1 Direct Methods .............................................. 72
  2.4.2 Discrete Methods ........................................... 72
2.4.3 Iterative Methods ............................................. 72
2.4.4 Multiresolution .............................................. 73
2.4.5 Summary ..................................................... 73

2.5 Evaluation Schemes ........................................... 74
2.5.1 Target Registration Error .................................... 74
2.5.2 Gold Standards .............................................. 74
2.5.3 Visual Inspection ............................................. 75
2.5.4 Consistency Registration Error ............................ 76
2.5.5 Image Matching Criteria ................................... 76
2.5.6 Clinical Usefulness .......................................... 76
2.5.7 Summary ..................................................... 77

2.6 Review of Registration Methods .............................. 77
2.6.1 Individual Methods .......................................... 77
2.6.2 Summary ..................................................... 85

2.7 Proposed Validation Method ................................... 87

2.8 Summary ..................................................... 89

3 Review of Biomechanical Breast Models ........................ 90
3.1 Introduction ..................................................... 90
3.2 Continuum Mechanics ......................................... 92
3.2.1 Kinematics .................................................. 92
3.2.2 Stress ....................................................... 95
3.2.3 Equilibrium Equations .................................... 97
3.2.4 Boundary Conditions ...................................... 97
3.2.5 Principle of Virtual Displacements ....................... 98
3.2.6 Summary ..................................................... 99

3.3 Material Models ................................................ 100
3.3.1 Elastic Materials ........................................... 100
3.3.2 Hyperelastic Materials .................................... 101
3.3.3 Summary ..................................................... 105

3.4 Finite Element Method ......................................... 105
3.4.1 Geometrical Discretization ................................ 105
3.4.2 Finite Element Formulation for Static Linear Elasticity .. 108
3.4.3 Finite Element Formulation for Static Non-Linear Elasticity . 111
### 3.4.4 Volume Preservation

3.4.5 Summary

### 3.5 Mechanical Breast Properties

### 3.6 Review of Biomechanical Breast Models

#### 3.6.1 Mesh Generation

#### 3.6.2 Boundary Conditions

#### 3.6.3 Tissue Properties

#### 3.6.4 Solution Strategies

#### 3.6.5 Evaluation

3.7 Summary

### 4 Review of MR Breast Lesion Segmentation and Classification

#### 4.1 Introduction

#### 4.2 Segmentation

##### 4.2.1 Introduction

##### 4.2.2 Overview of Segmentation Methods

##### 4.2.3 Review of DCE MR Mammography Segmentation

##### 4.2.4 Summary

#### 4.3 Features

##### 4.3.1 Introduction

##### 4.3.2 Feature Candidates

##### 4.3.3 Unsupervised Feature Selection

##### 4.3.4 Summary

#### 4.4 Classification

##### 4.4.1 Introduction

##### 4.4.2 Linear Discriminant Analysis

##### 4.4.3 Logistic Regression

##### 4.4.4 Evaluation Measures and Techniques

##### 4.4.5 Summary

#### 4.5 Summary

### 5 Comparison of Biomechanical Breast Models

#### 5.1 Introduction

#### 5.2 Materials and Methods

##### 5.2.1 Datasets
5.2.2 Non-rigid Registration .................................. 151
5.2.3 Biomechanical Breast Models ......................... 152
5.2.4 Landmark Selection .................................. 160
5.2.5 Error Measures ....................................... 161
5.2.6 Overview of Tests .................................... 164
5.2.7 Summary ............................................... 165

5.3 Results ...................................................... 166
5.3.1 Visualisation of Results ................................ 166
5.3.2 Test1: Direct Solvers .................................. 166
5.3.3 Test2: Mesh Density ................................... 166
5.3.4 Test3: Finite Deformation Formulation (Volunteer 1, BC1 only) .... 170
5.3.5 Test4: Linear and Piecewise-Linear Models, Skin, Poisson’s Ratios, Boundary Conditions ....................... 171
5.3.6 Test5: Transverse Isotropic Models .................. 176

5.4 Summary ...................................................... 178

6 Validation of DCE MR Mammography Registration .. 180
6.1 Introduction .................................................. 180
6.2 Material and Methods ..................................... 181
6.2.1 Datasets ................................................. 181
6.2.2 Biomechanical Breast Models ......................... 181
6.2.3 Simulation of Breast Deformations .................... 182
6.2.4 Registration .............................................. 182
6.2.5 Error Quantification ................................... 184
6.2.6 Selecting the Better Registration Strategies ......... 185
6.2.7 Validation Strategy ..................................... 186
6.2.8 Sensitivity and Significance of Results ............... 186
6.2.9 Analysis of Registration Error ......................... 187

6.3 Results ...................................................... 187
6.3.1 Selecting the Better Registration Strategies ......... 187
6.3.2 Validation on Test Data ................................ 190
6.3.3 Sensitivity and Significance of Results ............... 190
6.3.4 Analysis of Registration Error ......................... 193
6.3.5 Computational Costs ................................... 196
# Contents

6.4 Summary ................................................. 196

7 Computer Aided Diagnosis System for DCE MR Mammography .................................................. 200
  7.1 Introduction ............................................. 200
  7.2 Materials and Methods ................................. 201
    7.2.1 Datasets ............................................. 201
    7.2.2 Image Registration .................................. 201
    7.2.3 Segmentation Refinement ......................... 202
    7.2.4 Features ............................................. 205
    7.2.5 Classification ....................................... 206
  7.3 Results ................................................. 207
    7.3.1 Segmentation ....................................... 207
    7.3.2 Feature Extraction .................................. 211
    7.3.3 Classification ....................................... 213
    7.3.4 Individual Feature Performance .................. 213
    7.3.5 Multivariate Classification ....................... 221
  7.4 Summary ................................................. 226

8 Conclusions and Future Work ................................................. 228
  8.1 Summary ................................................. 228
    8.1.1 Biomechanical Breast Models ...................... 228
    8.1.2 Validation of Registration ......................... 229
    8.1.3 CAD System .......................................... 229
  8.2 Conclusions and Discussion .............................. 230
  8.3 Future Work ............................................. 231
    8.3.1 CAD System .......................................... 231
    8.3.2 Biomechanical Breast Models ...................... 232
    8.3.3 Validation of Registration ......................... 232

Bibliography ................................................. 233

Publications ................................................. 258
List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Breast anatomy</td>
<td>26</td>
</tr>
<tr>
<td>1.2</td>
<td>Microscopic changes within the duct</td>
<td>27</td>
</tr>
<tr>
<td>1.3</td>
<td>Main MRI contrast mechanisms T1 and T2</td>
<td>28</td>
</tr>
<tr>
<td>1.4</td>
<td>Breast appearance on MRI</td>
<td>29</td>
</tr>
<tr>
<td>1.5</td>
<td>Illustration of DCE MR mammography acquisition</td>
<td>31</td>
</tr>
<tr>
<td>1.6</td>
<td>DCE MR mammography example</td>
<td>32</td>
</tr>
<tr>
<td>1.7</td>
<td>Terms for relative intensity-time curve pattern</td>
<td>33</td>
</tr>
<tr>
<td>1.8</td>
<td>Terms for morphological image features</td>
<td>34</td>
</tr>
<tr>
<td>1.9</td>
<td>Summary ROC curve</td>
<td>36</td>
</tr>
<tr>
<td>1.10</td>
<td>DCE MR mammography example with motion artifacts</td>
<td>37</td>
</tr>
<tr>
<td>1.11</td>
<td>Examples of substantial pectoral muscle movements</td>
<td>38</td>
</tr>
<tr>
<td>1.12</td>
<td>Thesis structure</td>
<td>40</td>
</tr>
<tr>
<td>2.1</td>
<td>Illustration of image registration</td>
<td>43</td>
</tr>
<tr>
<td>2.2</td>
<td>Example of a geometrical image transformation</td>
<td>44</td>
</tr>
<tr>
<td>2.3</td>
<td>Illustration of common registration approaches</td>
<td>45</td>
</tr>
<tr>
<td>2.4</td>
<td>The structure of the registration review chapter</td>
<td>46</td>
</tr>
<tr>
<td>2.5</td>
<td>2D deformation examples employing polynomial basis functions</td>
<td>51</td>
</tr>
<tr>
<td>2.6</td>
<td>2D deformation examples employing discrete cosine transform basis functions</td>
<td>52</td>
</tr>
<tr>
<td>2.7</td>
<td>2D deformation examples employing multiquadratic basis functions</td>
<td>53</td>
</tr>
<tr>
<td>2.8</td>
<td>2D deformation examples employing thin plate spline basis functions</td>
<td>54</td>
</tr>
<tr>
<td>2.9</td>
<td>2D deformation examples employing elastic body spline basis functions</td>
<td>55</td>
</tr>
<tr>
<td>2.10</td>
<td>2D deformation examples of cubic B-spline mesh and local linear interpolation</td>
<td>58</td>
</tr>
<tr>
<td>2.11</td>
<td>Example of extracting edges from DCE MR mammograms</td>
<td>66</td>
</tr>
<tr>
<td>2.12</td>
<td>Effects of motion and contrast-enhancement on the intensity relation of DCE MR mammograms</td>
<td>67</td>
</tr>
<tr>
<td>3.1</td>
<td>Outline of mechanical computer modelling process</td>
<td>90</td>
</tr>
<tr>
<td>3.2</td>
<td>The structure of the biomechanical breast model review chapter</td>
<td>91</td>
</tr>
<tr>
<td>3.3</td>
<td>Illustration of motion $\Phi$ acting on the continuum body $\mathcal{B}$</td>
<td>93</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>3.4</td>
<td>Forces and stress components of current configuration</td>
<td>96</td>
</tr>
<tr>
<td>3.5</td>
<td>Deformation tests for incompressible, isotropic hyperelastic materials.</td>
<td>103</td>
</tr>
<tr>
<td>3.6</td>
<td>Fitting of hyperelastic material models</td>
<td>104</td>
</tr>
<tr>
<td>3.7</td>
<td>FEM domain discretization</td>
<td>105</td>
</tr>
<tr>
<td>3.8</td>
<td>Tetrahedron configuration</td>
<td>106</td>
</tr>
<tr>
<td>3.9</td>
<td>Stress-strain relationships for models from the literature</td>
<td>120</td>
</tr>
<tr>
<td>3.10</td>
<td>Relative stress-strain relationships for models from the literature</td>
<td>120</td>
</tr>
<tr>
<td>4.1</td>
<td>Computer aided diagnosis system for DCE MR mammography</td>
<td>124</td>
</tr>
<tr>
<td>5.1</td>
<td>Geometric models for volunteer one</td>
<td>153</td>
</tr>
<tr>
<td>5.2</td>
<td>Geometric models for volunteer two</td>
<td>154</td>
</tr>
<tr>
<td>5.3</td>
<td>Stress-strain relationships of fatty tissue for MM1-MM9</td>
<td>157</td>
</tr>
<tr>
<td>5.4</td>
<td>Stress-strain relationships of fibroglandular tissue for MM1-MM9</td>
<td>157</td>
</tr>
<tr>
<td>5.5</td>
<td>Relative stress-strain relationships for MM1-MM9</td>
<td>157</td>
</tr>
<tr>
<td>5.6</td>
<td>Stress-strain relationships of fatty tissue for hyperelastic models</td>
<td>158</td>
</tr>
<tr>
<td>5.7</td>
<td>Stress-strain relationships of fibroglandular tissue for hyperelastic models</td>
<td>158</td>
</tr>
<tr>
<td>5.8</td>
<td>Relative stress-strain relationships for hyperelastic models</td>
<td>158</td>
</tr>
<tr>
<td>5.9</td>
<td>Examples of corresponding landmarks for volunteer one</td>
<td>162</td>
</tr>
<tr>
<td>5.10</td>
<td>Examples of corresponding landmarks for volunteer two</td>
<td>163</td>
</tr>
<tr>
<td>5.11</td>
<td>Landmark distribution and displacement in the deformed breast</td>
<td>164</td>
</tr>
<tr>
<td>5.12</td>
<td>In-vivo breast deformation for volunteer one</td>
<td>167</td>
</tr>
<tr>
<td>5.13</td>
<td>In-vivo breast deformation for volunteer two</td>
<td>168</td>
</tr>
<tr>
<td>5.14</td>
<td>Displacement errors for volunteer one, BC1, MM1-MM9</td>
<td>173</td>
</tr>
<tr>
<td>5.15</td>
<td>Displacement errors for volunteer one, BC2, MM1-MM9</td>
<td>173</td>
</tr>
<tr>
<td>5.16</td>
<td>Displacement errors for volunteer one, BC3, MM1-MM9</td>
<td>174</td>
</tr>
<tr>
<td>5.17</td>
<td>Displacement errors for volunteer two, BC1, MM1-MM9</td>
<td>174</td>
</tr>
<tr>
<td>5.18</td>
<td>Displacement errors for volunteer two, BC2, MM1-MM9</td>
<td>175</td>
</tr>
<tr>
<td>5.19</td>
<td>Displacement errors for volunteer two, BC3, MM1-MM9</td>
<td>175</td>
</tr>
<tr>
<td>5.20</td>
<td>Displacement errors of transverse isotropic material for volunteer one</td>
<td>177</td>
</tr>
<tr>
<td>5.21</td>
<td>Displacement errors of transverse isotropic material for volunteer two</td>
<td>177</td>
</tr>
<tr>
<td>6.1</td>
<td>Breast surface displacement due to tension of pectoral muscle</td>
<td>183</td>
</tr>
<tr>
<td>6.2</td>
<td>Maximum breast surface displacement after non-rigid registration</td>
<td>183</td>
</tr>
<tr>
<td>6.3</td>
<td>Influence of image similarity measure on registration performance</td>
<td>188</td>
</tr>
</tbody>
</table>
6.4 Registration performance for training set ........................................ 189
6.5 Examples images from training set ................................................. 191
6.6 Examples images from test set .................................................... 192
6.7 Influence of image enhancement on registration performance ............ 194
6.8 Example images of three types of source images ............................. 195
7.1 Example of initial crude outline ..................................................... 208
7.2 Result of principle component analysis .......................................... 208
7.3 Example of class conditional probability density functions ................. 208
7.4 Example of maximum a posteriori segmentation ............................. 208
7.5 Segmentation example for original image sequence .......................... 209
7.6 Segmentation example after registration ........................................ 209
7.7 Mean overlap between refined and manual segmentation .................. 210
7.8 Clustering of features ............................................................... 211
7.9 Influence of registration and segmentation refinement on the t-test probability ......................................................... 215
7.10 Influence of registration and segmentation refinement on the leave-one-out AUC ................................................................. 215
7.11 Empirical ROC curves for the 11 least correlated features ............... 222
7.12 Smooth ROC curves for the 11 least correlated features .................. 223
7.13 Change of feature values due to registration .................................. 225
7.14 Example of maximum washout decrease ..................................... 226
7.15 Example of signal enhancement ratio increase ................................ 226
List of Tables

1.1 Definitions of morphological image feature terms .......................... 34
1.2 Prospective DCE MR mammography studies published in 1993 to 2000 ... 35

2.1 Juxtaposition of global basis functions ........................................ 56
2.2 Mean regularization costs associated with basis functions ................. 62
2.3 Euler-Lagrange equations for differential quadratic forms ................. 63
2.4 Automatic registration algorithms for DCE MR mammography ............ 86
2.5 Non-automatic registration algorithms for DCE MR mammography ........ 87

3.1 Basic properties of common material model ................................... 100
3.2 Reported Young’s moduli for fatty, fibroglandular and tumourous tissue ... 117
3.3 Mesh configurations of reported biomechanical breast models .............. 119
3.4 Young’s moduli of breast tissues for models from the literature ............ 121

4.1 Overview of computerized classification studies for MR breast lesions .... 131
4.2 Classification performance of shape features from previous studies ......... 133
4.3 Classification performance of margin features from previous studies ....... 135
4.4 Classification performance of kinetic features from previous studies ...... 138
4.5 Classification performance of texture features from previous studies ...... 140
4.6 Overview of promising features ........................................... 142
4.7 Classification table .................................................. 146

5.1 Geometric FEM Model Configuration ........................................... 152
5.2 Material properties of linear and piecewise-linear models .................... 156
5.3 Material properties of hyperelastic models .................................... 156
5.4 Configurations of boundary conditions ....................................... 159
5.5 Overview of test configurations ............................................. 165
5.6 Change of results due to different FEM solvers and geometric models ...... 169
5.7 Finite deformation results .................................................. 170
5.8 Result of best model configurations ......................................... 176
5.9 Sensitivity analysis of the displacement error .................................. 176
<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.1</td>
<td>Overview of registration configurations</td>
<td>184</td>
</tr>
<tr>
<td>6.2</td>
<td>Sensitivity analysis of registration performance</td>
<td>193</td>
</tr>
<tr>
<td>7.1</td>
<td>List of assessed registration strategies</td>
<td>202</td>
</tr>
<tr>
<td>7.2</td>
<td>List of assessed segmentation strategies</td>
<td>204</td>
</tr>
<tr>
<td>7.3</td>
<td>MARIBS scoring system</td>
<td>206</td>
</tr>
<tr>
<td>7.4</td>
<td>Ranking of segmentation quality</td>
<td>211</td>
</tr>
<tr>
<td>7.5</td>
<td>Overview of the 11 least correlated features</td>
<td>212</td>
</tr>
<tr>
<td>7.6</td>
<td>Description of classifiers</td>
<td>213</td>
</tr>
<tr>
<td>7.7</td>
<td>T-test probability of individual features</td>
<td>216</td>
</tr>
<tr>
<td>7.8</td>
<td>Leave-one-out area under the ROC curve for individual features</td>
<td>217</td>
</tr>
<tr>
<td>7.9</td>
<td>Characteristics of shape, margin and other features</td>
<td>218</td>
</tr>
<tr>
<td>7.10</td>
<td>Characteristics of kinetic features</td>
<td>219</td>
</tr>
<tr>
<td>7.11</td>
<td>Characteristics of kinetic and texture features</td>
<td>220</td>
</tr>
<tr>
<td>7.12</td>
<td>Leave-one-out classification result</td>
<td>221</td>
</tr>
<tr>
<td>7.13</td>
<td>Probabilities that the area under smooth ROC curve is the same for two methods</td>
<td>224</td>
</tr>
<tr>
<td>7.14</td>
<td>Leave-one-out classification result for perturbed initial outlines</td>
<td>224</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

Worldwide, breast cancer is by far the most common cancer in women. One in nine women in the UK will develop breast cancer at some point in their lives [CRUK, 2004]. Relative to other cancers, it has a favourable prognosis with a mortality to incidence ratio of about 36% worldwide. Still, it remains the leading cause of cancer death in women [Parkin, 2000]. This thesis presents novel contributions that ultimately aim to improve breast cancer diagnosis.

1.1 Breast Cancer Detection

Worldwide, X-ray mammography is currently the accepted standard for screening the breast. Currently in the UK breast screening is offered every 3 years to all women aged between 50 and 70, and on request to women older than that. However, prior to 2004 the age limit was 64. Ten years after the introduction of the NHS breast screening program in 1988, the mortality rate has fallen by 20%, while the incidence rate increased by about 15% [National Statistics, 2003]. The results of eight randomized trials support that this improvement originates from the early detection of breast cancer rather than improved treatment. Meta-analyses of these trials showed that breast cancer mortality of the screened group decreased by 20% to 35% [Fletcher and Elmore, 2003].

X-ray mammography has, however, some well-known limitations. Health risk associated with the X-ray radiation do impose restrictions on the quality and quantity of images that can be obtained. Depending on a number of factors, 4% to 34% of cancers are missed [Huynh et al., 1998], while 5% to 10% of biopsies\(^1\) are benign\(^2\) [Fletcher and Elmore, 2003; Carney et al., 2003; Barlow et al., 2004]. Especially dense\(^3\) breast tissue, which is more common in young women, causes a statistically significant reduction in sensitivity\(^4\) and specificity\(^5\) [Carney et al., 2003; Barlow et al., 2004]. While the incidence of breast cancer below the age of 50 is low in the overall population (about 2%), women with a strong family history of breast

---

\(^1\)A biopsy is defined as the removal of a piece of living tissue for microscopic examination.

\(^2\)Benign tumours do not invade and destroy tissue or spread to distant sides in the body.

\(^3\)Glandular breast tissue (described in Section 1.2) is known as dense breast tissue.

\(^4\)Sensitivity measures the probability of detecting the disease when the disease exists.

\(^5\)Specificity measures the probability of identifying that there is no disease when the disease is absent.
cancer are known to develop breast cancer early [Claus et al., 1991]. About 5% to 10% of all breast cancer cases occur in women with a family history of breast cancer, and of these about 84% can be attributed to the gene mutations BRCA1 and BRCA2. By the age of 50, 45% of BRCA1 gene carriers and 28% of the BRCA2 gene carriers will have developed the disease [Ford et al., 1998]. Screening methods for young women who have a high-risk of developing breast cancer should be improved.

The main alternative imaging technologies currently used routinely as adjunct diagnostic tools for symptomatic patients are ultrasound and dynamic contrast-enhanced (DCE) magnetic resonance (MR) mammography. Ultrasound is most useful in distinguishing between solid masses and cysts; for investigating palpable abnormalities within dense tissue and for guiding core biopsies. Ultrasound is relatively inexpensive, can produce images in real-time and has no harmful effects. Drawbacks are its operator dependence, its restrictions in penetration depth and its poor image quality. DCE MR mammography provides information about tissue vascularity and permeability, which cannot be obtained by X-ray mammography. DCE MR mammography has been beneficial for the evaluation of equivocal X-ray mammographic findings, breast cancer staging, treatment planning and monitoring response to chemotherapy [Rankin, 2000; Kuhl and Schild, 2000; Ikeda et al., 2000]. DCE MR mammography is known to be safe, to have a very high sensitivity and unaffected by breast density. Its main disadvantages are lower specificity [Lawrence et al., 1998b; Hrung et al., 1999] and cost, i.e. DCE MR mammography can be 10 to 20 times more expensive than X-ray mammography [Lee and Weinreb, 2004].

Initial screening studies of high-risk women, which compared DCE MR mammography with X-ray mammography and ultrasound, reported improved sensitivity (96% versus 29% and 30%) at an intermediate specificity (94% versus 98% and 91%) [Kuhl et al., 2000; Warner et al., 2001; Podo et al., 2002]. Larger trials are currently underway in several countries, including the United States, Canada, the United Kingdom, Germany, the Netherlands, France and Italy. The success of DCE MR mammography for screening, where follow-up procedures are the largest cost components, is likely to depend on finding techniques to improve specificity without reducing sensitivity.

One hypothesis of this thesis is that reducing the errors introduced by patient motion will improve accuracy of DCE MR mammography. The following chapters provide back-

---

6Patients are symptomatic if they have the common symptoms associated with a disease. The most common symptom of breast cancer is a lump.

7A cyst is a closed cavity filled with fluid.
1.2 Breast Anatomy

The breast (or mamma) is a gland\(^8\) that produces milk (or lactate). Breasts are lifelong undeveloped and functionless in men, while in women they grow and differentiate at and after puberty. Under the influence of hormones, breasts change with the menstrual cycle in anticipation of pregnancy. They reach their full development at the final stage of pregnancy and during lactation. After the menopause they finally enter a resting state.

The adult female breast is essentially composed of four structures: glandular lobules, milk ducts, fat and connective tissue as depicted in Figure 1.1a. Glandular lobules and milk ducts are the important components for milk production. Each breast consists of 12 to 20 irregular shaped glandular lobes. Each lobe is drained by a separated milk duct with a microscopic opening on the nipple. Before reaching the nipple, each duct widens to form a cavity (ampulla) before it narrows for its opening. Each lobe is subdivided into numerous glandular lobules, which empty into the duct system of the lobe they serve. The lobules are the functional secretory units of the breast. Each lobule is composed of multiple cavities, called alveoli or acini. The alveoli and their feeding ducts, called intralobular terminal ducts, are lined by a continuous layer of epithelial cells, that are specialized in milk production. A discontinuous layer of myoepithelial cells lies under the layer of epithelial cells. In lactating women, the myoepithelial cells enable milk propulsion by contracting the alveoli and the ducts in response to the hormone oxytocin. The lobules and ducts are sealed and surrounded by an uninterrupted basement membrane.

The breast rests on a large fan-shaped muscle, the pectoralis major, as illustrated in Figure 1.1b. Originating from the front of the rib cage (from the clavicle to the 6th-7th rib), the fibres of the pectoralis major converge to the upper arm. The pectoralis major is a powerful muscle for drawing the arm forward across the chest and for rotating the arm around its centre. The pectoralis minor muscle lies underneath the pectoralis major and is used for lowering and rotation the shoulder blade. The breast is attached to the chest wall by fibrous strands, called Cooper’s ligaments. These strands emerge from the deep fascia\(^9\), spread throughout the entire breast and end on the skin. A stretching or softening of the

---

\(^8\)A gland is an organ or a group of cells that is specialized for synthesizing and secreting certain fluids.

\(^9\)A fascia is a sheet or band of fibrous connective tissue separating or binding together muscles and organs.
ligaments, which may occur during pregnancy or with old age, leads to a sagging of the breast in the upright position. Separation of the deep fascia and the pectoral fascia by an area of loose connective tissue, called retromammary space, allows the breast to move to some extent on the pectoral fascia. The lobes are held in place by fatty and loose connective tissue. The fatty tissue increases towards the periphery of the lobules. The nipple contains circularly arranged muscle fibers that respond to stimulation to make the nipple erect. Other muscles fibers are longitudinal and may retract the nipple. The nipple is surrounded by a circular pigmented area, the areola, which has many oil secreting glands for lubrication during lactation. The breast is covered by skin, which is usually thinner than the skin of the adjacent regions.

Like the rest of the body, the breast contains blood vessels, lymph vessels and nerves. The breast is supplied by multiple arteries as illustrated in Figure 1.1b. The arteries branch mostly in the region above the areola, causing the blood supply to the upper half of the breast to be almost twice that of the lower half. The blood is passed back to the heart via veins that closely follow the arteries. Breast tissue is drained by lymphatic vessels which mainly lead to lymph nodes in the armpit (or axilla) and behind the breastbone as shown in Figure 1.1b. Consequently, these are usually the first nodes involved when breast cancer
1.3 Breast Cancer

Most breast cancers arise in the epithelial component of the ducts and sometimes the alveoli of the lobules. About 80% of all invasive\(^\text{10}\) breast cancers have ductal involvement and about 10% have lobular components [Margolese et al., 2000]. Figure 1.2 illustrates common microscopic changes to the duct that may lead to invasive breast cancer. Growth of some extra cell linings, called ductal hyperplasia, increases the risk of developing breast cancer in the next 10 to 20 years by 1.8 times [Page and Dupont, 1993]. If these extra cells look strange, called atypical ductal hyperplasia, the risk is increased by 4-5 times. If the duct is filled with atypical cells, called ductal carcinoma in situ (DCIS), the risk is much higher, i.e. 8-10 times. Equivalent changes occur in the lobules and lead to similarly increased risks.

Further growth of invasive tumours leads to mass formation and distant spread. Tumour growth greater than a diameter of 1-2mm and spread to other sites generally depends on the growth of new blood vessels, called angiogenesis [Blood and Zetter, 1990]. Thus, tumours may be detected earlier on the basis of increased vessel density and vessel permeability. DCE MR

\(^{10}\)Invasive (or infiltrating) means that these cancers have spread into neighbouring tissue, i.e. have grown through the basement membrane.
mammography, described in the next section, is currently the most powerful image modality to show these effects.

1.4 DCE MR Mammography

1.4.1 MRI Basic Principles

MRI exploits the fact that materials with an odd number of protons or neutrons (e.g. hydrogen, fluorine-19, phosphorus-19) have a nuclear magnetic moment. When placed in a strong magnetic field, the normally randomly orientated nuclear moments align. This static magnetization is however too weak to be measured. Instead, radiofrequency (RF) pulses are used to tip the net magnetization away from the static field. The magnetization is then measured while it oscillates in a plane perpendicular to the static field. The two main MRI contrast mechanisms are called T1 and T2. T1 describes the rate at which the longitudinal magnetization (magnetization in direction of static field) is restored after it has been disturbed by RF pulses. As illustrated in Figure 1.3a, T1 contrast will be produced as long as images are acquired before full recovery has been achieved. T2 describes the rate at which the transverse magnetization (magnetization in the plane perpendicular to the static field) decays after it has been created. T2 contrast is produced by delaying the image acquisition (Figure 1.3b). Thus, MR image intensities are generally not an absolute measure but depend strongly on the employed pulse sequence. This makes MRI a very flexible image modality, that can provide information not only about proton density, but also about water diffusion and water content. Additionally information can be provided due to the different resonance frequencies of chemical elements.

Signals are spatially localized, i.e. images are formed, by applying a magnetic field gradient causing the magnetization in each region of the volume to oscillate at a distinct frequency.

![Figure 1.3: Illustration of main MRI contrast mechanisms T1 and T2.](image-url)
Figure 1.4: Illustration of the normal breast anatomy and its appearance on a high-resolution T1-weighted MR image. Tissue marked as fibroglandular includes milk ducts, glandular lobules and connective tissue as depicted in Figure 1.1a). The MR image has a coronal in-plane resolution of $0.66 \times 0.66$ mm and a slice thickness of 2.5 mm.

For 3D images, one can differentiate between two types of acquisition protocols, namely multislice and volume imaging. Individual slices are excited during multislice imaging, and localization within the slice is achieved by frequency encoding in one direction and phase encoding in the second direction. Volume imaging involves exciting the entire 3D region of interest and localization via frequency encoding in one direction and phase encoding in the other two directions. Volume imaging enables the acquisition of thin slices with excellent signal to noise ratios, but may require a longer acquisition time. Patient motion during volume imaging will effect the entire image while motion during multislice imaging will effect each slice differently. In both methods motion leads to blurring and ghosting.

Figure 1.4 shows annotated orthogonal slices of a 3D T1-weighted MR image of a breast. These images were acquired with the patient lying on her front (i.e. prone), her arms by her side and her breasts hanging under gravitational pull in the double breast coil\textsuperscript{11}. In this example and in general fatty and fibroglandular tissue can easily be discriminated due

\textsuperscript{11}A rounded rectangular box with insulated conductors wound around it used for receiving MR signals.
to their different T1 properties. Studies showed however, that the MR tissue parameters
did not provide sufficient information to reliably detect and diagnose breast cancer malign-
nancies [Heywang-Köbrunner and Beck, 1995; Kerslake et al., 1995]. Instead, studying the
blood distribution and dynamics within the breast, using dynamic contrast-enhanced MRI,
has proved to be clinically very useful. The rational for this method is the observation
that malignant tumours\textsuperscript{12} have generally an increased vascularity and an increased vascular
permeability compared with benign tissues [Blood and Zetter, 1990; Vaupel, 1994].

\subsection*{1.4.2 Contrast-Enhanced Image Sequence}

In DCE MR mammography, the signal from the blood is increased by intravenous adminis-
tration of gadolinium diethylenetriaminepentaacetic acid (Gd-DTPA). Gd-DTPA is the first
paramagnetic contrast agent approved for clinical use and shortens the rate at which the lon-
gitudinal magnetization is restored, i.e. T1. It distributes throughout the intravascular and
interstitial\textsuperscript{13} space [Weinmann et al., 1990] when injected into the bloodstream. Increased
vascularity, increased vascular permeability, and/or increased interstitial space (which are all
common signs of malignancies) lead to a faster distribution of the contrast agent into and
out of the interstitial space. The dynamics of the blood distribution is therefore observed
by taking MR images before and after the injection of the contrast agent as illustrated in
Figure 1.5. Acquisition of the so-called post-contrast images starts with the injection of the
contrast agent and images are continuously obtained for a time period of about 10 minutes.
Temporal resolution is therefore equivalent to the acquisition time of a single image.

Figure 1.6 shows example slices from a DCE MR mammography image sequence. In this
case, one pre-contrast and five post-contrast 3D T1-weighted images were acquired covering
both breasts with a field of view (FOV) of 340×170×160mm. The images have a spatial
resolution of 1.33×1.33×2.5mm and were obtained within 90s. This 3D image protocol is
typical for screening the breasts in search for malignancies. Improving the rather low temporal
resolution would require a decrease in the number of acquired image elements (called pixels in
2D and voxels in 3D). A coarser spatial resolution would however disadvantage the detection
of small tumours. A smaller FOV is an option when only a particular region has to be
investigated, like in the case of a single lesion with ambiguous features on the 3D DCE MR

\textsuperscript{12}Malignant tumours invade and destroy the tissue in which they originate and can spread to other sides
in the body via the bloodstream and the lymphatic system.

\textsuperscript{13}Small space in a tissue or between parts of the body, which is here equivalent to the extravascular and
extracellular space.
Figure 1.5: Illustration of DCE MR mammography acquisition for T1-weighted images. Observable effects due to contrast agent are generally a rapid enhancement (i.e. signal increase) of the heart and the blood vessels, and a slow steady enhancement of the fibroglandular tissue. Regions of fibroglandular tissue, with a rapid initial enhancement followed by an intensity decrease, are suspect malignancies.

mammograms. Acquiring just five 2D slices with 256×128 pixels each, for example, can be achieved within 12s [Brown et al., 2000].

1.4.3 Interpretation of DCE MR Mammography

The DCE MR mammograms are analysed by radiologists, who search for regions of strong and early enhancement. It is standard practise, that the analysis of this digital 4D dataset (3D images and time) is supported by subtraction images and relative time-intensity curves. Subtraction images, as shown on the right side of Figure 1.6, are produced by subtracting the pre-contrast image from the post-contrast images. The three relative time-intensity curves, as illustrated in the top right graph of Figure 1.6, show the change in mean intensity of user defined region of interest (ROIs) against time.

Benign and malignant lesions can often be successfully discriminated by comparing the dynamics of their blood distribution in this way. Benign lesions may, however, possess similar microvascular density [Buadu et al., 1996; Buckley et al., 1997] and vascular permeability as malignancies.

Discrimination relies then mainly on morphology. MR images often excellently display the growth pattern, the extent and the internal structure of the lesion. This can be useful to detect, for example, malignancies growing along ducts because of their ductlike shapes; infiltrating lesions due to their irregular shapes; malignancies with necrotic centres by their
Figure 1.6: Coronal example slices from a DCE MR mammography image sequence. The left side shows from top to bottom the pre-contrast image, and the 1st, 3rd and 5th post-contrast image. The right side shows from top to bottom the relative intensity-time curve for three manually selected regions of interest (ROIs) and the result after subtracting the pre-contrast image from the 1st, 3rd and 5th post-contrast image. The 1st subtraction image was annotated to show the location of the ROIs. The images have an in-plane resolution of 1.33×1.33mm and a slice thickness of 2.5mm.
rim enhancements; and fibroadenomas due to their well-defined margins and round shapes.

In summary, important image features of DCE MR mammography include the relative amount of enhancement; the pattern of the relative intensity-time curve; the lesion’s shape; the lesion’s margin; and the lesion’s spatial pattern of internal enhancement. The recommended terms for describing these image features are summarized in Figure 1.7 and 1.8 [Schnall and Ikeda, 1999; Ikeda et al., 2001]. Features which suggest malignancy include fast initial enhancement, washout, spiculated margin, irregular shape and heterogeneous rim enhancement. In contrast slow initial enhancement, persistent late enhancement, smooth margin, round shape and homogeneous enhancement are more likely to be signs of benign lesions [Abdolmaleki et al., 1997; Nunes et al., 1999; Brown et al., 2000; Kinkel et al., 2000]. The enhancing lesion in Figure 1.6 may be described as having a fast initial enhancement, a washout, an irregular shape, a spiculated margin and a heterogeneous enhancement, which would suggest that this is malignancy.

### 1.4.4 Prospective DCE MR Mammography Studies

Many studies have assessed the diagnostic power of DCE MR mammography. A meta-analysis performed by the author, showed that prospective studies with more than 40 symptomatic patients, published in the years 1993 to 2000, reported a sensitivity of 82.5% to 100% and a specificity of 36.7% to 86.4%, see Table 1.2. Their overall performance for all 1664 patients was 93.1% sensitivity and 68.5% specificity. Table 1.2 reveals the high variety in the patient cohort, the spatial and temporal resolution and the complexity of the classification rule for these studies. None of these study parameters correlated significantly with the classification accuracy. The highest accuracy was achieved by a 2D study with a very high

---

14 The overall sensitivity (specificity) was calculated by the number of all correctly classified malignant (benign) lesions divided by the number of all malignant (benign) lesions.
Figure 1.8: Terms for morphological image features from DCE MR mammography lexicon [Ikeda et al., 2001]. Further definitions are given in Table 1.1.

<table>
<thead>
<tr>
<th>Lesion type</th>
<th>Lesion subtype</th>
<th>Internal enhancement</th>
<th>Shape/margin</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foci</td>
<td></td>
<td></td>
<td>Smooth round</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smooth oval</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Lobulated</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Irregular</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Spiculated</td>
</tr>
<tr>
<td>Mass</td>
<td></td>
<td>Homogeneous confluent</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heterogeneous</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heterogeneous rim enhancement</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heterogeneous enhancing septations</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Heterogeneous dark septations</td>
<td></td>
</tr>
<tr>
<td>Non-mass-like enhancement</td>
<td>Linear, nonspecific</td>
<td>Homogeneous confluent</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Linear, ductal</td>
<td>Heterogeneous nonspecific</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Segmental</td>
<td>Heterogeneous stippled/punctate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Regional</td>
<td>Heterogeneous clumped</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diffuse, patchy</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Diffuse, nonspecific</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Asymmetric</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Smooth</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Irregular</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Clumped</td>
</tr>
</tbody>
</table>

Table 1.1: Definitions of morphological image feature terms from Figure 1.8.

temporal resolution, where the onset of the lesion’s enhancement was normalized to the time of the main artery’s (i.e. aortic) enhancement [Boetes et al., 1994]. Studies that derived their classification rule from previous retrospective studies (marked with * in Table 1.2) were amongst the better methods. Their overall specificity of 75% was however still relative low.
1.4 DCE MR Mammography

Table 1.2: Summary of prospective DCE MR mammography studies with more than 40 symptomatic patients that were published between 1993 and 2000 and provided sufficient information. Studies where images were interpreted according to the results from a previous retrospective study are marked with *.

<table>
<thead>
<tr>
<th>Reference</th>
<th>P</th>
<th>Age</th>
<th>M</th>
<th>D</th>
<th>VV</th>
<th>Δt</th>
<th>C</th>
<th>Sen</th>
<th>Sp</th>
<th>Acc</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Cross et al., 1993]</td>
<td>41</td>
<td>68</td>
<td>6</td>
<td></td>
<td>0.7</td>
<td>300</td>
<td>1</td>
<td>95.3</td>
<td>36.7</td>
<td>66.0</td>
</tr>
<tr>
<td>[Gilles et al., 1996]</td>
<td>172</td>
<td>54</td>
<td>47</td>
<td>69</td>
<td>3.5</td>
<td>47</td>
<td>2</td>
<td>95.0</td>
<td>51.1</td>
<td>73.0</td>
</tr>
<tr>
<td>[Obdeijn et al., 1996]</td>
<td>54</td>
<td>61</td>
<td>12</td>
<td></td>
<td>7.2</td>
<td>120</td>
<td>2</td>
<td>90.9</td>
<td>66.7</td>
<td>78.8</td>
</tr>
<tr>
<td>[Fischer et al., 1999b]</td>
<td>463</td>
<td>54</td>
<td>74</td>
<td>12</td>
<td>6.3</td>
<td>87</td>
<td>8</td>
<td>92.6</td>
<td>65.0</td>
<td>78.8</td>
</tr>
<tr>
<td>[Müller-Schimpfe et al., 1997]</td>
<td>89</td>
<td>53</td>
<td>60</td>
<td>2</td>
<td></td>
<td>9.6</td>
<td>60</td>
<td>6</td>
<td>94.9</td>
<td>64.1</td>
</tr>
<tr>
<td>[Fobben et al., 1995]</td>
<td>89</td>
<td>23</td>
<td>19</td>
<td></td>
<td>5.3</td>
<td>135</td>
<td>3</td>
<td>82.5</td>
<td>79.1</td>
<td>80.8</td>
</tr>
<tr>
<td>[Abdolmaleki et al., 1997]*</td>
<td>56</td>
<td>51</td>
<td>73</td>
<td>2</td>
<td>5.5</td>
<td>13</td>
<td>10</td>
<td>97.6</td>
<td>66.7</td>
<td>82.2</td>
</tr>
<tr>
<td>[Bonè et al., 1997]</td>
<td>220</td>
<td>52</td>
<td>61</td>
<td>10</td>
<td></td>
<td>5.5</td>
<td>360</td>
<td>3</td>
<td>92.4</td>
<td>72.0</td>
</tr>
<tr>
<td>[Kvistad et al., 2000]</td>
<td>130</td>
<td>53</td>
<td>55</td>
<td>6</td>
<td>7.6</td>
<td>57</td>
<td>2</td>
<td>87.5</td>
<td>79.3</td>
<td>83.4</td>
</tr>
<tr>
<td>[Kacl et al., 1998]</td>
<td>50</td>
<td>57</td>
<td>60</td>
<td>13</td>
<td>8.6</td>
<td>60</td>
<td>6</td>
<td>92.1</td>
<td>76.0</td>
<td>84.0</td>
</tr>
<tr>
<td>[Nunes et al., 1997a]*</td>
<td>94</td>
<td>50</td>
<td>49</td>
<td>9</td>
<td>0.5</td>
<td>155</td>
<td>7</td>
<td>100.0</td>
<td>68.8</td>
<td>84.4</td>
</tr>
<tr>
<td>[Kinkel et al., 2000]*</td>
<td>57</td>
<td>52</td>
<td>60</td>
<td>9</td>
<td>1.3</td>
<td>300</td>
<td>5</td>
<td>91.2</td>
<td>82.6</td>
<td>86.9</td>
</tr>
<tr>
<td>[Helbich et al., 1997]</td>
<td>66</td>
<td>47</td>
<td>35</td>
<td>4</td>
<td>4.4</td>
<td>87</td>
<td>2</td>
<td>96.2</td>
<td>81.6</td>
<td>88.9</td>
</tr>
<tr>
<td>[Boetes et al., 1994]*</td>
<td>83</td>
<td>51</td>
<td>75</td>
<td>9</td>
<td>37.4</td>
<td>2</td>
<td>3</td>
<td>95.4</td>
<td>86.4</td>
<td>90.9</td>
</tr>
<tr>
<td>Overall</td>
<td>1664</td>
<td>53</td>
<td>61</td>
<td>14</td>
<td>7.0</td>
<td>128</td>
<td>5</td>
<td>93.1</td>
<td>68.5</td>
<td>80.8</td>
</tr>
<tr>
<td>Overall*</td>
<td>290</td>
<td>51</td>
<td>63</td>
<td>8</td>
<td>12.2</td>
<td>112</td>
<td>6</td>
<td>96.3</td>
<td>75.0</td>
<td>85.6</td>
</tr>
</tbody>
</table>

The performance of several diagnostic tests can be represented by a summary receiver operating characteristic (ROC) curve [Walter, 2002]. A ROC curve shows how the sensitivity and the specificity changes when different thresholds are applied to the classification rule. Figure 1.9 depicts the smooth summary ROC curves for the prospective DCE MR mammography studies from Table 1.2. These were created by the author using a regression model as described in [Walter, 2002]. The summary ROC curves show that studies with very low specificity could have been the result of a classification rule that strongly favoured a high
sensitivity. The area under the ROC curve (AUC) is a common measure to summarize the classification performance. The 4 studies, that derived their classification rule from previous retrospective studies and are marked with *, achieved a slightly better AUC of 0.92 than all 11 studies together. These summary ROC curves predict that a 95% sensitivity will result in a 65% (78%) specificity for the 11 (4) studies.

Further advances in acquisition technology and interpretation rules will certainly help to improve the classification performance. One error source of DCE MR mammography which is often overlooked is patient motion. For example, only two of the studies from Table 1.2 have taken some measures to reduce patient motion [Boetes et al., 1994; Nunes et al., 1997a]. As image resolution improves and quantitative analysis becomes more sophisticated, small movements that previously could have been ignored will have a greater effect.

1.5 Patient Motion and Registration

Patient motion during DCE MR mammography hinders the detection and diagnosis of contrast-enhanced lesions. Radiologists are faced with the problem of distinguishing intensity changes caused by contrast agent from intensity changes caused by patient motion. This is illustrated in Figure 1.10, where the relative time-intensity curves are quite similar.
although ROI2 and ROI3 contain malignant tissue (infiltrating ductal carcinoma) while ROI1 and ROI4 enclose benign tissue (according to 6 years of follow-up). Also note the cardiac motion artifacts at the bottom of the images in Figure 1.10. These artifacts make the interpretation of both axillas very difficult. Slight motion may not impair the diagnosis of large lesions if radiologists tediously study the original images. But for small lesions, detectability is limited and quantitative evaluation becomes useless. In the case of substantial motion, reading becomes impossible and the acquisition of DCE MR mammograms has to be repeated after at least 7 hours of the previous acquisition to allow for normal clearance of Gd-DTPA.
Spatial image alignment, i.e. where corresponding image features are at the same spatial position, will especially benefit radiologists’ assessment as well as computer aided detection and diagnosis methods for smaller lesions.

Some measures can be taken to prevent motion artifacts. Cardiac motion artifacts within the breast or the axillas can be avoided by choosing a coronal or sagittal image plane. Respiratory motion artifacts can be decreased by imaging the patient in the prone position. Patient movements commonly observed include relaxation and contraction of the pectoral muscle as shown in Figure 1.11, and repositioning of the chest or arms. While some fixation of the breast during acquisition is certainly helpful, a strong compression of the breast would alter the dynamics of the contrast-agent uptake and hence is not advisable. Furthermore, this fixation would only slightly reduce local artifacts introduced by pectoral muscle movement, which is a frequently observed phenomenon. From experience, patient motion during DCE MR mammography is very common. More than 90% of all DCE MR mammography image sequences have visually recognizable motion artifacts (excluding cardiac motion artifacts on axial images).
Previously, an algorithm has been devised for the non-rigid registration\textsuperscript{15} of images [Rueckert \textit{et al.}, 1999b] and applied to DCE MR mammograms. It was shown that this registration method significantly improved the quality of the subtraction images as visually judged by radiologists [Denton \textit{et al.}, 1999]. A first quantitative assessment of the registration results showed, however, volume changes of up to 20\% for the region of the enhanced lesions [Tanner \textit{et al.}, 2000]. Such volume changes are highly unlikely during a single DCE MR mammography acquisition since no significant change in external force is applied to the breast and all images are acquired within 10 minutes. While the registration can be constrained to preserve volume, the problem is still how to measure the residual registration error since the true patient movement (termed here the ground truth) is unknown. Evaluation of the performance of DCE MR mammography registration algorithms has mainly relied on visual inspection, which is impossible in regions of contrast enhancement. This work is addressing this unsatisfactory state of validation. The first objective of this thesis is therefore to develop a technique for the validation of registration algorithms for correcting for patient motion in DCE MR mammography.

\section*{1.6 Computer Aided Diagnosis System}

Reading DCE MR mammograms is time consuming and difficult because the whole image sequence has to be examined for kinetic and morphological patterns of malignancies. However, the challenge is even greater for screening, where there is a higher percentage of smaller lesions. Computerized systems, that can reduce the workload of the radiologists, are an attractive proposition considering the worldwide shortage of radiologists [The Royal College of Radiologists, 2002].

Computer aided detection and diagnosis systems aim to remove the need for having the images read by two independent radiologists. These systems alert the radiologist to the fact that a malignant growth may be present and provide associated classification suggestions. In contrast, computer aided diagnosis (CAD) systems aim to solve the vital subtask of supplying classification suggestions for marked image regions.

Several CAD systems have been proposed for DCE MR mammography, including those outlined in [Sinha \textit{et al.}, 1997; Gilhuijs \textit{et al.}, 1998; Gilhuijs \textit{et al.}, 2002; Sonoda, 2003; Gibbs and Turnbull, 2003; Chen \textit{et al.}, 2004]. Their standalone classification performance,\textsuperscript{15} A rigid registration can in the best case remove all artifacts that were caused by the movement of a rigid body (i.e. global translation and rotation). Non-rigid methods also account for more flexible movements like deformations. More details about registration methods can be found in Chapter 2.
as measured by the area under the ROC curve, ranges from 0.80 to 0.96. This is promising when compared to the radiologists’ performance of 0.92 as estimated by the summary ROC curve, see Figure 1.9. However, some of these results could be too optimistic, since more feature candidates than recommended were assessed when training the classifiers [Hair et al., 1998]. Furthermore, only one study [Sonoda, 2003] used registered images. The second objective of this thesis is therefore to improve the classification of breast lesions by creating a CAD system for DCE MR mammography which includes image registration.

1.7 Thesis Structure

Figure 1.12 provides an overview of the chapters in this thesis. The individual chapters are then described.

Chapter 1 is the present introductory chapter.

In Chapter 2, known registration methods for DCE MR mammography are reviewed with regard to their transformation models, matching criteria, optimization strategies and evaluation schemes. Evaluation is identified as inadequate and a novel validation method is proposed. This validation method is based on simulating gold standard deformations by biomechanical models using finite element methods (FEMs).

An overview of continuum mechanics, material models and finite element methods is given in Chapter 3. This chapter then ends with a discussion of the mechanical properties of breasts and a review of published biomechanical breast models.

Figure 1.12: The structure of this thesis. The experiments conducted by the author are described in Chapter 5 to Chapter 7 and the white arrows show the dependencies between these chapters.
Chapter 4 reviews previous automatic methods for segmenting and classifying MR breast lesions. The aim of this survey was to find a suitable framework for the computer aided diagnosis system which is described in Chapter 7. Promising image features are also identified and defined in this chapter.

The accuracy of biomechanical breast models to predict the displacement of internal breast structures for true deformations is evaluated in Chapter 5.

Chapter 6 describes the validation of the non-rigid registration for DCE MR mammography with respect to registration accuracy. The validation was based on simulating plausible breast deformations using a suitable biomechanical breast model as determined in Chapter 5.

The developed computer aided diagnosis system for DCE MR mammography is presented in Chapter 7. This chapter describes the classification performance and the influence of registration on the results.

Finally, the eighth and last chapter is a summary and discussion of the work presented in this thesis, and outlines areas for further work.

1.8 Contributions of this Research

The main contributions of the work described in this thesis are:

- A comprehensive review of DCE MR mammography registration techniques including their evaluation strategies (Chapter 2);

- The first quantitative analysis of the impact of non-rigid registration for motion correction on the MR breast lesion volume [Tanner et al., 2000];

- Development of a novel validation method for non-rigid registration algorithms for DCE MR mammography (Chapter 5 and 6);

- A systematic assessment of factors influencing the accuracy of biomechanical breast models, including the use of transverse isotropic models for breast deformations for the first time (Chapter 5);

- An extensive quantitative validation of one non-rigid registration algorithm for the registration of dynamic contrast-enhanced MR breast images (Chapter 6);

- Development and evaluation of a computer aided diagnosis system for DCE MR mammography including registration and a novel semi-automatic segmentation method (Chapter 7);
1.9 Software

The image registration software validated in this thesis was written by Daniel Rueckert and Julia Schnabel.

FEM modelling was conducted using the commercial package ANSYS [ANSYS Inc., 2000], version 5.7 and 7.1. A program for assigning material properties to individual elements was written by Andy Castellano-Smith. Software for image interpolation, target registration error calculation and error visualization was written by Julia Schnabel. Philipp Batchelor provided a program for the calculation of volume change for 10-noded tetrahedrons. The author produced programs for data-transfer, specification of FEM boundary conditions, assessment of landmark errors and creation of result summaries.

The segmentation refinement and classification algorithms were written by the author using MATLAB and its image processing and statistical toolboxes. Statistical significance of the difference between receiver operating characteristics curves was determined using the ROCKIT program from [Metz et al., 1998].
Chapter 2

Review of DCE MR Mammography Registration

This chapter reviews the current state of registering the image sequence acquired during dynamic contrast-enhanced (DCE) magnetic resonance (MR) mammography. The aim of this review is to explain the main components employed for DCE MR mammography registration, to provide an overview of existing methods and to discuss their strengths and weaknesses. Special attention was paid to the techniques used for the evaluation of the performance of these methods and a common weakness amongst all published work was identified. This leads to a novel validation method, which is also proposed in this chapter. The development, evaluation and application of this novel validation method is described in subsequent chapters of this thesis.

2.1 Introduction

The process of determining the spatial correspondence between features in images or real world objects is known as registration. One application of registration is to align images, such that corresponding image features are at the same spatial position. This process is illustrated in Figure 2.1. A and B are two images of the same head. Apart from differences

![Diagram of image registration](image)

Figure 2.1: Illustration of image registration. The images $A$ and $B$ can be aligned by sampling the source image $B$ at a clockwise rotation of $25^\circ$ and a translation.
in intensity, the images are different because the head in image B is tilted and shifted. When superimposing the images (overlay(A,B)) one can observe that for example the white nose is not at the same location as the black nose, i.e. the noses are not aligned. One can also identify some image positions which should correspond, like the tip of the nose. Such recognizable image features are called landmarks. Having identified corresponding landmarks, one can display their spatial relationship by displacement vectors as shown by the black arrows in the top right image (correspondence(A,B)). These vectors point from the landmark position \( x \) in A to the corresponding landmark position \( T(x) \) in B. If such a vector is known for every voxel position \( x \) in image A (the reference image), then image B (the source image) can be geometrically transformed to align with A by sampling B at \( T(x) \) as shown in Figure 2.2. The optimal transformation \( T \) for Figure 2.1 is a clockwise rotation of 25° and a translation.

Having illustrated the principle concept of image registration, the question remains how this task could be achieved automatically. Figure 2.3 outlines two common approaches for image registration. Key components include the transformation model, the image matching criterion and the optimization strategy.
Figure 2.3: Illustration of two common registration approaches. The image focused approach (left) starts by matching independently local image features (e.g. edges, local intensity statistics). A continuous global transformation $T$ is then deduced by fitting a transformation model to the established correspondences. In the transformation focused approach (right), the transformation parameters are optimized with respect to a global image matching criterion (e.g. total edge overlap, global intensity statistics).

- The transformation model defines the space of admissible transformations. It can be used to prevent implausible transformations (i.e. incorporate prior knowledge), to restrict the number of possible solutions or to define the interpolation function between corresponding points. The transformation model may be based on a set of parameters (e.g. rotation angles) or on an optimization criterion (e.g. minimize volume change).

- The image matching criterion is a surrogate measure for estimating the unknown error of alignment from the image content. This measure may be based on the difference of extracted corresponding features or on the expected relationship between the intensities of corresponding voxels. Ideally, this measure should have a single global optimum when the images are in spatial correspondence. Noise, acquisition artifacts, the limited information provided by the images and the often unknown relationship between the image pairs make it generally difficult to establish such an ideal image matching criterion.

- The optimization strategy defines how the space of admissible transformations is explored. Time restrictions, the number of free transformation parameters and the com-
putational cost for calculating the image matching criterion generally prohibit an ex-
hauative evaluation of the whole search space for non-rigid transformations. Assuming
a well-behaved cost function, standard iterative optimization techniques like Gradient
Descent, Newton method or Conjugate Gradient Descent may be applied instead [Press
et al., 1992].

A large number of image registration algorithms have been proposed. General reviews
and surveys of these can be found in [Gottesfeld-Brown, 1992; van den Elsen et al., 1993;
Maintz and Viergever, 1998; Rohr, 2000; Hill et al., 2001; Mäkelä et al., 2002; Pluim et al.,
2003; Zitová and Flusser, 2003; Kostelec and Periaswamy, 2003]. Registration methods for
DCE MR mammography have been reported since 1995. These methods are reviewed in
this chapter as shown in Figure 2.4. Concepts and definitions for DCE MR mammography
registration are provided in Sections 2.2 to 2.4. These sections give details of the employed
transformation models, the image matching criteria and the optimization strategies. Another
important aspect is the evaluation scheme used to assess the registration outcome. These
schemes are examined in Section 2.5. Each registration algorithm is reviewed individually
in Section 2.6 and all registration methods are juxtaposed in Table 2.4 and Table 2.5 at the
end of Section 2.6. A new validation method is finally proposed in Section 2.7. The chapter
ends with a summary of the review (Section 2.8).

Figure 2.4: The structure of this chapter.
2.2 Transformation Models

Prior knowledge about plausible transformations can be incorporated into the registration via the transformation model. A model which only permits 3D shifts (or translations) and 3D rotations is for example very suitable for registering 3D images of the same rigid object. Such a global rigid transformation is, however, an inaccurate motion model for most of the human body with its complex composition of deformable soft tissue, independently moving local rigid structures and sliding tissue interfaces. Modelling motion of these structures requires transformations with more degrees of freedom (or unknowns). Prior knowledge about the mechanical properties of the imaged organ could be used to constrain these transformations in general. Modelling complex compositions will, however, require the identification of individual structures.

Suitable transformation models are most difficult to identify when the registration problem is not to recover organ motion, but to assess anatomical differences over time or between subjects. Newly appearing structures or anatomical variations preclude the use of, for example, transformations which are volume preserving or continuous. Even the definition of spatial correspondence is debatable in this context.

Let us now return to the reviewed application. MR images are acquired while the patient is lying on her front (known as prone) with her breasts hanging in the breast coil. The breast coil may be padded with soft material to reduce motion of the breasts. Depending on the imaging protocol, the arms are either folded around the head or placed alongside the body. As described in Section 1.2, breasts have a very heterogeneous composition and deform easily. Patient movements commonly affecting DCE MR mammography include flexion and relaxation of the pectoral muscles as shown in Figure 1.11, and repositioning of the chest or arms. Cardiac motion artifacts within the breast and the axillas can be avoided by choosing a coronal or sagittal image plane. Respiratory motion artifacts are decreased since patients are imaged in the prone position. The registration of DCE MR mammography images acquired during a single visit is therefore a problem of recovering the breasts’ movements.

In the following, the various transformation models applied to DCE MR mammography will be described and illustrated. These illustrations are based on experiments conducted by the author. In each case, transformation $T(x)$ maps position $x = [x_1 \ x_2 \ x_3]^T$ in image $A$ to the corresponding position $T(x) = [T_1(x) \ T_2(x) \ T_3(x)]^T = x + u(x)$ in image $B$, where $u(x) = [u_1(x) \ u_2(x) \ u_3(x)]^T$ is the displacement vector at $x$. 
2.2 Transformation Models

Global transformation models describe the correspondence function by a relatively small number of parameters which affect the spatial transformation of all image positions. These models include linear, global polynomial and harmonic transformations. Each model is now introduced in turn.

2.2.1 Linear Transformation

What is generally (and also in this thesis) called a linear transformation is actually an affine transformation.

**Affine Transformation**

The 3D affine transformation consists of a (truly) 3D linear transformation followed by a 3D translation

\[
T_{\text{affine}}(x) = \begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix},
\]

where \(a_1, a_2, ..., a_9\) are the parameters of the (truly) 3D linear transformation and \(\tau_1, \tau_2, \tau_3\) are the 3 translation parameters. The (truly) linear transformation is a subclass of the affine transformation because the translation vector violates the linear relationship: \(T(\alpha x^{(1)} + \beta x^{(2)}) = \alpha T(x^{(1)}) + \beta T(x^{(2)})\) for all \(\alpha\) and \(\beta\). An affine transformation keeps lines straight and parallel. The affine transformation can be represented by translation, rotation, reflection, anisotropic scaling and shearing. Reflection is generally an unwanted transformation for medical image registration. Reflection can be avoided by rejecting (truly) linear transformations with a negative determinant or by optimizing only for translation, rotation, scaling and shearing. An important subclass of the affine transformation is the rigid transformation, which will be discussed next.

**Rigid Transformation**

A rigid transformation preserves the distance between points. The 3D rigid transformation \(T_{\text{rigid}}\) has 6 degrees of freedom: 3 translations \((\tau_1, \tau_2, \tau_3)\) and 3 rotation angles \((\alpha_1, \alpha_2, \alpha_3)\). A 3D rigid transformation which firstly rotates position \(x\) around the \(x_3\)-, \(x_2\)- and \(x_1\)-axis and then applies a translation is defined as

\[
T_{\text{rigid}}(x) = M_r x + [\tau_1 \ \tau_2 \ \tau_3]^T
\]

with rotation matrix \(M_r\)

\[
M_r = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha_1) & \sin(\alpha_1) \\ 0 & -\sin(\alpha_1) & \cos(\alpha_1) \end{bmatrix} \begin{bmatrix} \cos(\alpha_2) & 0 & -\sin(\alpha_2) \\ 0 & 1 & 0 \\ \sin(\alpha_2) & 0 & \cos(\alpha_2) \end{bmatrix} \begin{bmatrix} \cos(\alpha_3) & \sin(\alpha_3) & 0 \\ -\sin(\alpha_3) & \cos(\alpha_3) & 0 \\ 0 & 0 & 1 \end{bmatrix}
\]
where \( \alpha_1, \alpha_2 \) and \( \alpha_3 \) denote the rotation angles around the \( x_1 \), \( x_2 \) and \( x_3 \)-axis.

The motion of a 3D rigid object can be described by a 3D rigid transformation. It is therefore a very suitable transformation model for registering 3D images of individual bones or the head of a patient. The drawback of the rigid transformation over the affine transformation is rather small, since organs rarely just scale or shear. Affine registration is however useful to reduce scaling and skewing errors commonly introduced by the image acquisition process. Transformations with more degrees of freedom than the affine transformation are employed to compensate for more complicated deformations. These transformations are generally called non-rigid, non-linear or elastic. Non-rigid global transformations employed for DCE MR mammography are described next.

### 2.2.1.2 Global Basis Functions

If correspondence was only established for individual positions (e.g. displacement vectors for a set of landmarks), then the remaining displacements need to be inferred from these by interpolation. The selected type of interpolation function will restrict the possible transformations.

This problem can be formulated in the following way. Given a set of \( P \) points \( \mathbf{x}^{(p)} \) with associated displacement vectors \( \mathbf{u}^{(p)} \) for \( 1 \leq p \leq P \), determine a smooth and continuous transformation \( \mathbf{u}(\mathbf{x}) \) which minimizes the mean square error \( \sum_p |\mathbf{u}(\mathbf{x}^{(p)}) - \mathbf{u}^{(p)}|^2 \). Treating each coordinate independently, this can be viewed as fitting a smooth function \( \mathbf{u}_d(\mathbf{x}) \) to the points \( (\mathbf{x}^{(p)}, u_{dp}) \) for \( d = 1, 2, 3 \).

A very general and common choice for these functions is to define them by a linear combination of basis functions

\[
\mathbf{u}_d(\mathbf{x}) = \sum_{b=1}^{B} c_{db} h_b(\mathbf{x}) \quad \text{for } d = 1, 2, 3
\]

(2.3)

where \( B \) is the number of basis functions and \( c_{db} \) are unknown coefficients which are determined by solving a linear system of equations. This system of equations is derived from the points which should be interpolated and is given by

\[
\begin{bmatrix}
u_{d1} \\ u_{d2} \\ \vdots \\ u_{dP}
\end{bmatrix} =
\begin{bmatrix}
h_1(\mathbf{x}^{(1)}) & h_2(\mathbf{x}^{(1)}) & \cdots & h_B(\mathbf{x}^{(1)}) \\
h_1(\mathbf{x}^{(2)}) & h_2(\mathbf{x}^{(2)}) & \cdots & h_B(\mathbf{x}^{(2)}) \\
\vdots & \vdots & \ddots & \vdots \\
h_1(\mathbf{x}^{(P)}) & h_2(\mathbf{x}^{(P)}) & \cdots & h_B(\mathbf{x}^{(P)})
\end{bmatrix}
\begin{bmatrix}
c_{d1} \\ c_{d2} \\ \vdots \\ c_{dB}
\end{bmatrix}
\]

or

\[
\mathbf{u}_d = \mathbf{H} \mathbf{c}_d
\]

(2.4)

for \( d = 1, 2, 3 \).

For interpolation, matrix \( \mathbf{H} \) needs to be square \( (P = B) \) and non-singular. Equation (2.4) can then be solved by matrix inversion. \( \mathbf{H} \) will have more rows than columns \( (P > B) \) if there
are more established correspondences than basis functions, and (2.4) can only be solved in the least squares sense, i.e. minimize $\sum_d \sum_p \sum_b (c_{db} h_b - u_{dp})^2$. Established correspondences are then approximated rather than interpolated. Underdetermined systems, where $P < B$, require further constraining, like orthogonality$^1$ of the basis functions or volume preservation of the deformation. There are many numerical methods to solve a linear system of equations, see [Press et al., 1992]. In the special case of orthogonal basis functions, a faster and more accurate closed form solution to (2.4) exists and further basis functions can easily be added if required [Wolberg, 1990].

The choice of a basis function is non-trivial. Global basis functions which were used in published research as transformation models for DCE MR mammography registration are described next. They are also listed in Table 2.1 at the end of this section. Several of these basis functions can be defined in terms of the radial distance to the known data points, i.e. $h_p(x) = g(r^{(p)})$ with $r^{(p)} = |x - x^{(p)}| = \sqrt{\sum_d (x_d - x_{dp})^2}$. The interpolation behaviour of the discussed basis functions is compared using a simple 2D example with six landmark correspondences as shown in Figure 2.5. Assuming that the registration task is to account for motion of the breast during DCE MR mammography, more conservative interpolation functions (e.g. smooth, least straining, area preserving) are preferred.

**Polynomial** The basis functions of a 3D $K$th-order polynomial transformation model are defined by

$$x_1^l x_2^m x_3^n \quad \text{for} \quad l, m, n \geq 0 \quad \text{and} \quad l + m + n \leq K. \quad (2.5)$$

Note that the 1st-order polynomial transformation is equivalent to the affine transformation (2.1). Higher order polynomials are generally less useful because of their unpredictable behaviour [Gottesfeld-Brown, 1992]. Figure 2.5 shows a 2D example of interpolating six landmark correspondences by a 1st- or 2nd-order polynomial. The 2nd-order polynomial interpolated the correspondences, but produced a large deformation with a high area change.

$^1$Basis functions with an additional constraint $\sum_{p=1}^P h_i(x^{(p)}) h_j(x^{(p)}) = 0$ for $i \neq j$. 
Figure 2.5: 2D deformation examples employing polynomial basis functions. These examples were generated by the author. The left graph outlines the location of a group 20×20 pixels of the reference image and shows the landmark positions of the reference image (marked with •) and the source image (marked with ×). The four landmarks in the left lower corner are at the same location in both images, i.e. have zero displacement vectors. The other two landmarks have a horizontal displacement \( u_1 \) of +1 and -1 pixel. The middle and the right graph show the result of interpolating the six landmark displacements by a 1st- and 2nd-order polynomial. The grid in these images outlines the pixel locations at which the source image is sampled for generating a transformed source image. If a source landmark (×) does not lie on the intersection of dark lines in these images, then its displacement vector was only approximated. The basis function’s ability to produce area preserving transformations was quantified by the mean absolute area change of the transformation. The 1st-order polynomial, which is equivalent to an affine transformation, approximated the landmark displacements, preserved parallel lines and had a mean absolute area change of 6.0%. The 2nd-order polynomial interpolated the landmark displacements, produced a larger geometric distortion and was less area preserving (15.2%).

**Discrete Cosine Transform (DCT)**  
The 3D DCT with \( L \times M \times N \) lowest frequency basis functions is given by

\[
ud(x) = \sum_{l=0}^{L-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} B_l(i_1)B_m(i_2)B_n(i_3)c_{di,m,n} \quad \text{for} \quad d = 1, 2, 3 \quad (2.6)
\]

where \( i_d \) is the voxel index of \( x_d \) (see Notation), \( c_{di,m,n} \) are the coefficients of the DCT model and the \( m \)th DCT basis function is defined by

\[
B_m(i_d) = \alpha(m) \cos \left( \frac{(2i_d + 1)m\pi}{2I_d} \right)
\]

with \( \alpha(m) = \sqrt{1/I_d} \) for \( m = 0 \), \( \alpha(m) = \sqrt{2/I_d} \) for \( m = 1, \ldots, M - 1 \). The result of interpolating the previously described 2D example by a DCT deformation is shown in Figure 2.6. The DCT basis functions produce a large curved deformation with a high area change.
Figure 2.6: 2D deformation examples employing the discrete cosine transform (DCT) basis functions, which were produced by the author. Both DCT basis functions produced a large curved deformation with high mean absolute area changes of 15.9% (left) and 15.3% (right). Landmark displacements were approximated by employing four basis functions (2×2 frequencies) and interpolated otherwise.

**Multiquadratic** basis functions are defined by

\[ g(r) = \sqrt{r^2 + c} \]  

(2.7)

with locality parameter \( c > 0 \) and \( r = |\mathbf{x}| \). The locality parameter \( c \) controls the influence of distant landmark correspondences. As \( c \) shrinks, the region of influence of each landmark correspondence decreases. Common choices for \( c \) are the mean displacement of all landmarks or the mean distance between all reference landmarks. The multiquadratic basis functions produce small deformations with small area changes when interpolating the 2D example, see Figure 2.7.

**Thin Plate Spline (TPS)** was employed as a transformation model for DCE MR mammography registration by [Meyer et al., 1998]. The TPS is an exact interpolator which minimizes the so-called bending energy [Bookstein, 1989]. In 2D, the bending energy is defined by

\[ C_{BE_{2D}}(u) = \int_\Omega \left( \frac{\partial^2 u}{\partial x_1^2} \right)^2 + 2 \left( \frac{\partial^2 u}{\partial x_1 \partial x_2} \right)^2 + \left( \frac{\partial^2 u}{\partial x_2^2} \right)^2 \, d\mathbf{x} \]  

(2.8)

and it measures the strain energy of a thin plate of metal deformed by small vertical displacements \( u \). The bending energy for vector displacements is given by the sum of the bending energies of the individual components, i.e. \( C_{BE_{2D}}(\mathbf{u}) = \sum_d C_{BE_{2D}}(u_d) \). The basis function \( g(r) \) of the TPS is given by

\[ g(r) = r^2 \log r \quad (2D) \quad \text{and} \quad g(r) = r \quad (3D), \]  

(2.9)
Figure 2.7: 2D deformation examples employing multiquadratic basis functions ($r = |x|$). The examples were created by the author. The multiquadratic basis functions interpolated the landmark displacements and produced small deformations. Setting the locality parameter $c$ of the multiquadratic basis functions to either the mean landmark displacements ($c = 0.33$) or the mean distance of the reference landmarks ($c = 9.13$) produced very similar results with a low mean absolute area change of 6.0% and 6.8%, respectively.

where $r = |x|$. Equation (2.9) is the fundamental solution of the Euler-Lagrange equation associated with $C_{BE}$, i.e. the so-called biharmonic partial differential equation (PDE)

$$2
\nabla^4 u_i = \delta_x,$$

where $\delta_x$ is the Dirac distribution (see Section 2.2.3.2 for further details). Note that the 3D TPS basis function is equivalent to the multiquadratic basis function defined in (2.9) with $c = 0$. Figure 2.8 shows the results of employing TPS basis functions for the 2D example. For 3D basis functions, the example represents a case where all landmarks lie in the shown 2D plane and have only in-plane displacements ($u_{3}^{(p)} = 0$). The displacement field of any other 2D plane will be the same as the one shown. The 3D TPS produced a smaller deformation with less area change than the 2D TPS.

**Elastic Body Spline (EBS)** was proposed by [Davis et al., 1997] as a deformation model for MR breast image alignment. The EBS is based on the physical model of a homogeneous, isotropic 3D elastic body for small deformations. The 3D physical model is the system of Navier-Stokes equilibrium PDEs for linear elasticity, i.e.

$$\mu \nabla^2 u(x) + (\lambda + \mu) \nabla [\nabla^T \cdot u(x)] = -f(x)$$

or

$$\mu \sum_{j} \frac{\partial^2 u_d}{\partial x_j^2} + (\lambda + \mu) \frac{\partial}{\partial x_d} \sum_{j} \frac{\partial u_j}{\partial x_j} = -f_d(x) \quad \text{for} \quad d = 1, 2, 3$$

where $f = [f_1 \; f_2 \; f_3]^T$ is the force vector, $\lambda = E \nu / ((1 + \nu)(1 - 2\nu))$ and $\mu = E / (2(1 + \nu))$ are the Lamé constants, $E$ is Young’s modulus and $\nu$ is Poisson’s ratio [Timoshenko and
2.2 Transformation Models

Figure 2.8: 2D deformation examples employing thin plate spline (TPS) basis functions ($r = |x|$). The examples were produced by the author. The landmark displacements were interpolated by the 2D TPS and the 3D TPS. The 2D TPS created a large, curved deformation with a high mean absolute area change of 16.5%. The 3D TPS created a small deformation with a low area change (5.7%).

Goodier, 1970]. To avoid discontinuities from point forces, the force field was smoothed by

$$
f_{\alpha}(x) = \sum_{p=1}^{P} a^{(p)} r^{(p)} \quad \text{or} \quad f_{\beta}(x) = \sum_{p=1}^{P} a^{(p)} / r^{(p)}, \quad (2.11)$$

where $a^{(p)} = [a_{1p} a_{2p} a_{3p}]^T$ denotes the strength of the force field component of landmark $p$ with $r^{(p)} = |x - x^{(p)}|$. The basis functions of the EBS for the two smoothing functions in (2.11) are given by

$$
g_{\alpha}(r) = (\alpha r^2 I - 3xx^T)r \quad \text{or} \quad g_{\beta}(r) = \beta rI - xx^T / r, \quad (2.12)$$

respectively, where $\alpha = 12(1-v) - 1$, $\beta = 8(1-v) - 1$ and $I$ is the identity matrix. EBS$_{\alpha}$ and EBS$_{\beta}$ will be used to refer to these two functions. Of all global basis functions used for DCE MR mammography, the EBS was the only one where the cross effects of the displacements in the different directions were taken into account. The results of applying the EBS basis functions to the 2D example are shown in Figure 2.9. It can be observed that cross effects and Poisson’s ratio had a higher influence on EBS$_{\alpha}$ than on EBS$_{\beta}$. Smaller deformations with less area change were achieved by EBS$_{\beta}$. The lowest area change was obtained outside the physically meaningful range of Poisson’s ratio values, due to the neglectation of non-linear effects.
Figure 2.9: 2D deformation examples employing elastic body spline (EBS) basis functions \((r = |x|)\). The examples were created by the author. The top row (EBS\(_\alpha\)) and the bottom row (EBS\(_\beta\)) differ due to the employed smoothing function (see (2.11)). The deformations of the top row are more curved and much larger than the ones on the bottom row. EBS\(_\alpha\) and EBS\(_\beta\) reduce to the 3D volume spline and the 3D thin plate spline, respectively, if cross effects on \(u_2\) are ignored (left graphs with horizontal lines). Cross effects and Poisson’s ratio had a greater influence on EBS\(_\alpha\) than on EBS\(_\beta\). Landmark displacements were interpolated by all EBS basis functions. The mean absolute volume change for these deformations was 46.3%, 22.4%, 7.2%, 4.5% (top, left to right) and 5.7%, 4.6%, 2.9% and 0.5% (bottom, left to right) respectively.

### 2.2.1.3 Summary

This section discussed the global transformation models listed in Table 2.1. Using a 2D example with six corresponding landmarks, the interpolation behaviour of these models was illustrated. Great differences were observed with regard to deformation shape, deformation magnitude and local area preservation. Smallest deformations were achieved by the multi-quadratic, the 3D TPS and the EBS\(_\beta\) basis functions. Note that these functions are of similar form. Area was best preserved by the EBS\(_\beta\) with a Poisson’s ratio of 0.62.

The drawback of all global transformation models is that they influence all image positions, which makes optimization of their parameter computationally expensive. Models with parameters which have only a local influence will be discussed next.
### 2.2 Transformation Models

#### 2.2.2 Semi-Local Models

Semi-local models use a moderate number of parameters which influence only a part of the image. Thus, optimization of a single parameter is computationally less expensive than for global models. Cubic B-spline meshes and local linear interpolation methods fall into this category.

**Cubic B-Spline Meshes** were employed for image registration to interpolate or approximate the displacement vectors from a set of corresponding landmarks [Gottesfeld-Brown, 1992; Szeliski and Coughlan, 1994; Declerck et al., 1995]. [Rueckert et al., 1998] fully automated the B-spline based registration by replacing the manually selected corresponding landmarks by a regular mesh of pseudo landmarks (termed control points) whose displacement vectors are automatically optimized.

A 3D grid of approximating cubic B-splines is defined as follows. Let \( \{x \mid 0 \leq x_d \leq X_d, d = 1, 2, 3\} \) be the domain of the image volume and let the \( K_1 \times K_2 \times K_3 \) mesh of control points have spacing \( s_d = \frac{X_d}{K_d-1} \). Let the control point positions be \( [c_{1k_1,k_2,k_3} \ c_{2k_1,k_2,k_3} \ c_{3k_1,k_2,k_3}]^T \) for \( k_d \in \{0,1,\ldots,K_d-1\} \). \( T(x) \) is then determined from the position of the \( 4 \times 4 \times 4 \) control points.

<table>
<thead>
<tr>
<th>Name</th>
<th>Basis functions, ( h_b(x), b = 1, \ldots, B )</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D 1st-order polynomial</td>
<td>( 1, x_1, x_2, x_3 )</td>
<td>affine</td>
</tr>
<tr>
<td>3D ( K )-th order polynomial</td>
<td>( x_1^l x_2^m x_3^n )</td>
<td>( l, m, n \geq 0; l + m + n \leq K )</td>
</tr>
<tr>
<td>3D discrete cosine transform</td>
<td>( h_{l+mL+nLM+1}(x) = B_l(i_1)B_m(i_2)B_n(i_3) )</td>
<td>( l = 0, \ldots, L - 1 )</td>
</tr>
<tr>
<td></td>
<td>( B_m(i_d) = \alpha(m) \cos \left( \frac{(2i_d+1)m\pi}{2L} \right) )</td>
<td>( m = 0, \ldots, M - 1 )</td>
</tr>
<tr>
<td></td>
<td>( \alpha(m) = \sqrt{1/I_d} ) for ( m = 0 )</td>
<td>( n = 0, \ldots, N - 1 )</td>
</tr>
<tr>
<td></td>
<td>( \alpha(m) = \sqrt{2/I_d} ) otherwise</td>
<td>( I_d, i_d ) as defined</td>
</tr>
<tr>
<td>Radial basis functions</td>
<td>( h_b(x) = g(r^{(b)}) ) with ( r^{(b)} =</td>
<td>x - x^{(b)}</td>
</tr>
<tr>
<td>Multiquadratic</td>
<td>( \sqrt{r^2 + c} )</td>
<td></td>
</tr>
<tr>
<td>2D thin plate spline</td>
<td>( r^2 \log r )</td>
<td></td>
</tr>
<tr>
<td>3D thin plate spline</td>
<td>( r )</td>
<td></td>
</tr>
<tr>
<td>3D volume spline</td>
<td>( r^3 )</td>
<td></td>
</tr>
<tr>
<td>3D elastic body spline (EBS ( \alpha ))</td>
<td>( [\alpha r^2 I - 3xx^T]r )</td>
<td>( \alpha = 12(1 - v) - 1 )</td>
</tr>
<tr>
<td>3D elastic body spline (EBS ( \beta ))</td>
<td>( \beta r I - xx^T/r )</td>
<td>( \beta = 8(1 - v) - 1 )</td>
</tr>
</tbody>
</table>

Table 2.1: Juxtaposition of global basis functions used as transformation models for DCE MR mammography registration.
points surrounding $x$ by

$$T_{BSpline_d}(x) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(t_1)B_m(t_2)B_n(t_3)c_{d_1+l,k_2+m,k_3+n} \quad \text{for} \quad d = 1, 2, 3 \quad (2.13)$$

where $k_d = [\frac{d}{s_d}] - 1$, $t_d = \frac{d}{s_d} - [\frac{d}{s_d}] \in [0, 1]$ and $B_m$ is the $m$-th basis function of the B-spline defined as [Lee et al., 1997]:

\begin{align*}
B_0(t) &= (1-t)^3/6 & B_2(t) &= (-3t^3 + 3t^2 + 3t + 1)/6 \\
B_1(t) &= (3t^3 - 6t^2 + 4)/6 & B_3(t) &= t^3/6.
\end{align*}

The 2D example from Section 2.2.1 cannot be interpolated by a cubic B-spline mesh, since the problem is underdetermined (only 6 landmark correspondences but at least 16 unknown control point positions). Instead the behaviour of the cubic B-spline functions is illustrated by interpolating some displacement vectors from the solutions of global basis functions as shown in Figure 2.10. Interpolation by cubic B-splines achieves very similar results as the global basis functions. This demonstrates the flexibility of cubic B-spline meshes to model very different transformations as long as a sufficient number of control points are available.

**Local Linear Interpolation** refers to methods where intermediate displacement vectors are derived from neighbouring established displacement vectors by linear interpolation. Figure 2.10 shows the results of linearly interpolating regularly spaced displacement vectors from the solutions of global basis functions. Curved lines are clearly approximated by straight lines and discontinuous gradients at the established displacement vectors are common.

### 2.2.3 Local Models

Local, nonparametric models allow the most general transformations, i.e. with the most degrees of freedom, and have therefore the largest and most complex search space. The final vector field is not represented by a set of parameters, but usually by a displacement vector for each voxel position. This huge search space and the limited and noisy information provided by the images makes registration an ill-posed problem, i.e. existence, uniqueness and stability of a solution can not be guaranteed.

Through regularization, ill-posed problems may be reformulated into well-posed problems of variational calculus, where the problem is formulated as minimizing a scalar cost function [Thikonov and Arsenin, 1977]. This cost function is based on weighting the data fit (e.g. image dissimilarity) against the regularizing term (e.g. roughness of the transformation). Integration over the region of interest provides then a scalar function.

A system of PDEs emerges from variational forms as one derives equations necessary for
Figure 2.10: 2D deformation examples of cubic B-spline and local linear interpolation, which were generated by the author. The top row shows the solutions of interpolating the 2D example by three global basis functions as described in Section 2.2.1. The other rows show the results of interpolating some displacement vectors from the top row solutions by cubic B-splines (middle) and by local linear functions (bottom). These displacement vectors are located at the initial control point positions (●) and end at the positions marked by ×. As expected, curved lines are more accurately interpolated by cubic B-splines than by linear interpolation. Discontinuous gradients are removed by cubic B-splines and introduced by linear interpolation.
a minimum\textsuperscript{2} or tries to locate local or global minima by the gradient descent method. PDEs may also be derived directly from physical models like linear elasticity or viscous fluid flow.

The search space can also be restricted implicitly by smoothing the transformation after optimization with respect to the image matching criterion. The convergence behaviour for this method is generally not well understood.

This section reviews the local transformation models employed for DCE MR mammography alignment. Regularization via a scalar cost function (i.e. variational form) is most common and is described first. The relationship between variational forms and PDEs is then outlined and variational forms are compared. Finally, local models based on implicit smoothing are described.

### 2.2.3.1 Variational Forms

Variational forms are based on minimization of a scalar cost function \( C \), which generally consists of two terms, i.e.

\[
C(T) = C_{\text{mismatch}}(A, B^T) + \lambda C_{\text{regularization}}(T)
\]  

(2.14)

where \( C_{\text{mismatch}}(A, B^T) \) estimates the error of misalignment as derived from the image \( A \) and the transformed image \( B^T \) (e.g. image dissimilarity or distance between corresponding landmarks, see Section 2.3). \( C_{\text{regularization}} \) measures the deviation of the transformation from the assumed optimality (e.g. bending energy or volume change) and \( \lambda \) is a constant weighting the importance of the two terms. The optimal transformation \( T^* \) is then given by

\[
T^* = \arg\min_{T \in T} C(T)
\]

where \( T \) is the set of admissible transformations.

Common regularization functions are first-order differential quadratic forms, i.e.

\[
C_{\text{regularization}}(T) = \int_{\Omega T} \phi(\nabla T(x))dx = \int_{\Omega T} \sum_{i,j,k,l} a_{i,j,k,l} \frac{\partial T_i}{\partial x_j} \frac{\partial T_l}{\partial x_k} dx
\]  

(2.15)

or second-order differential quadratic forms, i.e.

\[
C_{\text{regularization}}(T) = \int_{\Omega T} \phi(\mathcal{H} T(x))dx
\]

\[
= \int_{\Omega T} \sum_{i,j,k,l,m,n} a_{i,j,k,l,m,n} \frac{\partial^2 T_i}{\partial x_j \partial x_k} \frac{\partial^2 T_l}{\partial x_m \partial x_n} dx
\]

(2.16)

where \( \nabla T \) denotes the Jacobian matrix of \( T \), \( \mathcal{H} T \) is the Hessian of \( T \) and \( \phi \) represents a quadratic function of the individual elements. The regularization function may also be expressed with respect to the displacement vector \( u(x) \). Note that \( \nabla T(x) = \nabla u(x) + I \) and \( \mathcal{H} T(x) = \mathcal{H} u(x) \).

\textsuperscript{2}The first derivative of the cost function is zero at an extremum.
These regularizers may be derived from physical models. Minimization of strain energy falls into this category. Strain energy is the energy stored in an elastically deformed body and is equivalent to the work done to deform the body. For linear materials, the strain energy $W$ (defined per unit volume) is given by

$$W = \frac{1}{2} \sigma : \epsilon,$$

where $\sigma$ is the stress tensor and $\epsilon$ is the strain tensor with elements $\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} + \sum_k \frac{\partial u_k}{\partial x_i} \frac{\partial u_k}{\partial x_j} \right)$. The last term can be neglected for small displacements \cite{Timoshenko1970}, i.e. $\epsilon_{ij} = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$.

Based on the aforementioned concept, several regularization terms have been proposed.

**Membrane Energy (ME)** is based on the strain energy of a membrane. The strain energy of a homogeneous, flexible membrane, which is stretched horizontally and deformed by a small vertical deflection $u$ is given by

$$C_{ME2D}(u) = \int_{\Omega_T} ||\nabla u||^2 \, dx = \int_{\Omega_T} \left[ \left( \frac{\partial u}{\partial x_1} \right)^2 + \left( \frac{\partial u}{\partial x_2} \right)^2 \right] \, dx. \quad (2.17)$$

The membrane energy for vector displacements is defined as the sum of the membrane energies of the individual components, i.e. $C_{ME2D}(u) = \sum_d C_{ME2D}(u_d)$. For 3D this results in

$$C_{ME3D}(u) = \int_{\Omega_T} \sum_{i=1}^{3} \sum_{j=1}^{3} \left( \frac{\partial u_i}{\partial x_j} \right)^2 \, dx. \quad (2.18)$$

Although this model has a physical meaning, it does not reflect the expected breast deformations.

**Linear-Elastic Energy** is based on the strain energy of a linear isotropic elastic material. Assuming small displacements, the stress-strain relationship of linear isotropic materials is given by $\sigma = \lambda \sum_i \epsilon_{ii} I + 2\mu \epsilon$ where $\lambda = \frac{\nu E}{(1+\nu)(1-2\nu)}$ and $\mu = \frac{E}{2(1+\nu)}$ are the Lamé constants defined with respect to Young’s modulus $E$ and Poisson’s ratio $\nu$ and $I$ is the identity matrix \cite{Timoshenko1970}. The regularization term is then given by

$$C_{\text{elastic}}(u) = \int_{\Omega_T} W \, dx = \int_{\Omega_T} \left[ \frac{\mu}{2} \sum_i \sum_j \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2 + \frac{\lambda}{2} \left( \sum_j \frac{\partial u_j}{\partial x_j} \right)^2 \right] \, dx. \quad (2.19)$$

Note that the second term of Equation (2.19) penalizes non-zero divergence, i.e. promotes volume preservation\footnote{Volume preserving transformations have a zero divergence. Divergence is however not equivalent to volume change, because $\sum_{d=1}^{3} \frac{\partial u_d}{\partial x_d} = \sum_{d=1}^{3} (\lambda_d - 1)$ whereas volume change is given by $\det(\nabla T) - 1 = \lambda_1\lambda_2\lambda_3 - 1$, where $\lambda_d$ denote the eigenvalues of $\nabla T$.}. $C_{\text{elastic}}$ was pioneered for image registration by \cite{Broit1981} and \cite{Bajcsy1989}. While this regularization term does not account for non-linear effects (non-linear materials and large deformations) and heterogeneous materials, it approximates at least an appropriate measure for the expected deformations.
Volume Preservation  The local volume change at position $x$ after applying transformation $T$ can be calculated by the determinant of the Jacobian matrix, i.e. $\det(\nabla T(x))$ where $(\nabla T(x))_{ij} = \frac{\partial T_i}{\partial x_j}$. $\det(\nabla T(x))$ provides the ratio of the transformed volume to the initial volume at $x$. The volume is therefore preserved at $x$ if

$$\det(\nabla T(x)) = 1. \quad (2.20)$$

For the cubic B-spline mesh transformation model (2.13), [Rohlfing et al., 2003] used the mean absolute logarithm of $\det(\nabla T(x))$ at the control point positions as a regularization term for volume preservation, i.e.

$$C_{volume}(T) = \frac{1}{K_1K_2K_3} \sum_{k_1=0}^{K_1-1} \sum_{k_2=0}^{K_2-1} \sum_{k_3=0}^{K_3-1} \left| \ln(\det(\nabla T([k_1s_1 k_2s_2 k_3s_3]^T))) \right|, \quad (2.21)$$

where a mesh of $K_1 \times K_2 \times K_3$ of control points with spacing $s_1 \times s_2 \times s_3$ was employed. Note that volume shrinkage and expansion are equally penalized by (2.21). Volume preservation is a meaningful constraint for many deformations of the human body, since tissue is generally incompressible. Exceptions include fluid filled spaces such as the heart, the lungs, the bladder, and deformations created by applying large forces (X-ray mammography).

Bending Energy (BE) refers to the strain energy of a thin metal plate subject to only slight bending (2.8). Its 3D equivalent is given by

$$C_{BE_{3D}}(u) = \sum_d \int_{\Omega^T} \left[ \sum_{j=1}^{3} \left( \frac{\partial^2 u_d}{\partial x_j^2} \right)^2 + \sum_{j=1}^{3} \sum_{k=j+1}^{3} 2 \left( \frac{\partial^2 u_d}{\partial x_j \partial x_k} \right)^2 \right] dx. \quad (2.22)$$

Bending energy penalizes changes in the gradient of the deformation field.

Curvature Energy is defined as

$$C_{CE}(u) = \sum_d \int_{\Omega^T} \left( \sum_j \frac{\partial^2 u_d}{\partial x_j^2} \right)^2 dx. \quad (2.23)$$

The integral terms approximate the curvature of the displacement component. Curvature energy and bending energy are very similar and have identical gradients (see Section 2.2.3.2).

Table 2.2 lists the regularization costs associated with the 2D deformation examples shown in Figures 2.5 to 2.9. It can be observed that bending energy and curvature energy led to similar costs. Depending on the regularization function, the minimum cost was achieved by either the 3D thin plate spline, a 3D elastic body spline or a multiquadratic basis function.
Table 2.2: Mean regularization costs associated with the 2D deformation examples created by the author and shown in Figures 2.5 to 2.9. The regularization functions were membrane energy (ME), linear-elastic energy (LE) using $v = 0.495$ and $E = 1$, volume preservation (VP), bending energy (BE) and curvature energy (CE) as defined in Equation (2.17) to (2.23). The lowest cost for each regularization function is marked by a box.

<table>
<thead>
<tr>
<th>Example transformation</th>
<th>Mean regularization cost (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Basis function</td>
</tr>
<tr>
<td>Polynomial</td>
<td>2nd-order</td>
</tr>
<tr>
<td>Discrete cosine transform</td>
<td>2x2 freq.</td>
</tr>
<tr>
<td>Discrete cosine transform</td>
<td>2x3 freq.</td>
</tr>
<tr>
<td>Multiquadratic</td>
<td>sqrt($r^2+0.33$)</td>
</tr>
<tr>
<td>Multiquadratic</td>
<td>sqrt($r^2+9.13$)</td>
</tr>
<tr>
<td>2D thin plate spline</td>
<td>$r^2 \log r$</td>
</tr>
<tr>
<td>3D thin plate spline</td>
<td>$r^3$</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\alpha}$, v=0</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\alpha}$, v=0.5</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\alpha}$, v=0.6</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\beta}$, v=0</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\beta}$, v=0.5</td>
</tr>
<tr>
<td>3D elastic body spline</td>
<td>EBS$_{\beta}$, v=0.62</td>
</tr>
</tbody>
</table>

2.2.3.2 Differential Forms

A system of PDEs may arise from the scalar cost functions $C(T)$ (defined in Section 2.2.3.1) when the optimization technique depends on the calculation of gradients.

The first variation of $C(T)$ at $T \in T$ in the direction $k \in T$, termed the Gâteaux derivative, is defined by [Borowski and Borwein, 1989]

$$\delta C(T, k) = \lim_{\epsilon \to 0} \frac{C(T + \epsilon k) - C(T)}{\epsilon} = \nabla C(T).$$  \hspace{1cm} (2.24)

For a minimizer $T^*$, $\delta C(T^*, k) = 0$ must hold for every $k \in T$. The equations $\delta C(T, k) = 0$ are called the Euler-Lagrange equations associated with $C$.

Equation (2.24) also defines the gradient that should be followed in a gradient descent solution approach, i.e. starting from an initial estimate $T_0$, the next transformation is then given by $T_{k+1} = T_k - \nabla C(T_k)$ for $k \geq 0$.

From calculus of variations it is known that cost functions of the form

$$C(u) = \int_{a}^{b} L(x, u(x), u'(x))dx$$  \hspace{1cm} (2.25)
Table 2.3: Euler-Lagrange equations for first- and second-order isotropic differential quadratic forms [Cachier and Ayache, 2004].

have the Euler-Lagrange equation

$$\delta C(u) = \frac{\partial L}{\partial u} - \frac{d}{dx} \left( \frac{\partial L}{\partial u'} \right) = \frac{\partial^2 L}{\partial u \partial u'} - \frac{\partial^2 L}{\partial u \partial u'} \frac{\partial u}{\partial x} \frac{\partial^2 u}{\partial x^2} = 0 \quad (2.26)$$

where $u' = \partial u / \partial x$ [Weisstein, 2003]. Two integrals may give rise to the same Euler-Lagrange equation, but will have different boundary conditions. Generally, it cannot be guaranteed that integral (2.25) has an extremum.

Table 2.3 lists the Euler-Lagrange equations for the components of first-order isotropic differential quadratic forms (2.15), i.e. $L = \phi(\nabla u(x))$. A cost function $C(u)$ with this regularization term has therefore Euler-Lagrange equations

$$\alpha \nabla^2 u_d(x) + \beta \nabla \nabla^T \cdot u(x) = -f_d(x) \quad \text{for} \quad d = 1, 2, 3, \quad (2.27)$$

where the force $f_d(x)$ is the derivative of the image dissimilarity $C_{\text{mismatch}}$ with respect to $u_d$. For the linear elastic regularization model $C_{\text{elastic}}$ (2.19), the values of $\alpha$ and $\beta$ in (2.27) are given by $\mu$ and $\mu + \lambda$, where $\mu$ and $\lambda$ are the Lamé constants describing the elastic material properties as defined in (2.19).

Similarly, second-order isotropic differential quadratic forms (2.16) have Euler-Lagrange equations of the form

$$\alpha \nabla^4 u_d(x) + \beta \nabla^2 \nabla \nabla^T \cdot u(x) = -f_d(x) \quad \text{for} \quad d = 1, 2, 3. \quad (2.28)$$

Regularizations using bending energy $C_{BE}$ (2.22) or curvature energy $C_{CE}$ (2.23) have the same Euler-Lagrange equations, namely $\alpha \nabla^4 u_d(x) = -f_d(x)$. Their boundary conditions are however different with $C_{BE}$ imposing tighter constraints [Brady and Horn, 1983]. Their fundamental solution is given by the thin plate spline basis function (2.9).

---

4 A condition imposed on the solution of a differential equation to obtain the desired particular solution.
Of the employed regularization terms, only the linear-elastic model $C_{\text{elastic}}$ (2.19) and the volume preserving regularization $C_{\text{volume}}$ (2.21) enables cross-effects between the components of the vector field. Otherwise ($C_{\text{ME}}, C_{\text{BE}}, C_{\text{CE}}$), regularization is applied independently to the individual displacement components $(u_1, u_2, u_3)$.

### 2.2.3.3 Smoothing

Implicit regularization by smoothing the deformation field was based either on local averaging or on Gaussian filtering. This method is guaranteed to produce smoother transformation, i.e. reduce the curvature of the transformation field. Other regularization criteria are not necessarily improved. Smoothing does for example not reduce the volume change introduced by a global scaling. Its global application causes sufficiently smooth regions to be unnecessarily smoothed. The effects of this method are generally not well understood.

### 2.2.3.4 Summary

Local transformation models for DCE MR mammography registration were most commonly defined by regularization functions. Only two of these functions, namely linear-elastic energy and volume preservation, took cross-effects into account. The diversity of the five regularization functions was illustrated by calculating the costs induced by fourteen 2D deformation examples. Depending on the regularization function, the minimum cost was yielded by either the 3D thin plate spline, a 3D elastic body spline or a multiquadratic basis function. The costs for bending energy and curvature energy were very similar. The regularization cost varied greatly otherwise, making it hard to choose an adequate factor for weighting it against the image matching criteria.

### 2.2.4 Summary

This section described global, semi-local and local transformation models employed for DCE MR mammography registration. Global basis functions are generally useful for deriving a dense deformation field from a sparse set of known correspondences. The diversity of their interpolations was illustrated. Semi-local models have the advantage of influencing only a part of the image and hence are computationally less expensive than global functions for the same number of parameters. Cubic B-spline models were better than local linear interpolation in modelling very different transformations. Local models are the most flexible transformations and are commonly defined by regularization functions. It is however difficult to select an adequate factor for weighting the regularization cost against the image matching criteria.

Several global and local models were inspired by mechanical models. Of these only a few (3D elastic body spline, linear-elastic energy, volume preservation) were based on 3D
2.3 Image Matching Criteria

This section describes the image matching criteria employed for registering DCE MR mammograms. The image matching criteria should measure the degree of image alignment. Depending on the information extracted from the images, the image matching criteria can be categorized into two types, namely feature and voxel based similarity measures.

The aim in feature based methods is to extract image characteristics which are anatomically meaningful and can reliably be detected in both images. Registration is then based on identifying corresponding features in the images and fitting a transformation model to the point correspondences.

In voxel based methods the quality of alignment is directly derived from the image intensities of corresponding voxels by defining a measure of similarity. The transformation parameters are then optimized to maximize similarity.

Feature and voxel based similarity measures which were employed for registering DCE MR mammograms are described in the following subsections.

2.3.1 Feature Based Similarity Measures

Most algorithms for registering DCE MR mammograms relied on voxel based similarity measures except [Huwer et al., 1996; Lucht et al., 2000; Srikanthana et al., 2002].

[Huwer et al., 1996] used a weighted sum of various image features to determine the pixel with the most similar image region, i.e.

\[ C_{Huwer} = w_g \Delta_g + w_{lmg} \Delta_{lmg} + w_{spat} \Delta_{spat} + w_{norm} \Delta_{norm} + w_{angle} \Delta_{angle} \]  \hspace{1cm} (2.29)

where \( \Delta_g \) is the grey-level difference, \( \Delta_{lmg} \) is the local mean grey-level difference, \( \Delta_{spat} \) denotes the Euclidean distance between the position of the input pattern and the currently estimated position, \( \Delta_{norm} \) is the gradient norm distance, \( \Delta_{angle} \) is the angle between two gradient vectors and finally \( w_p \) are experimentally determined weights. This measure assumes preservation of image intensities and image edges, which does not hold in enhancing regions. Establishing weights which are generally applicable is not trivial.

[Lucht et al., 2000] chose to extract local gray-level changes (edges) as features. The 2D images were filtered with an edge preserving median filter (9x9 pixels) to reduce noise, a gradient operator (of unspecified type) was applied, and a morphological opening (using
2.3 Image Matching Criteria

a 5 pixel cross) was performed. Edges of 1 to 2 pixel thickness were then extracted by thresholding the difference of the opened image and the gradient image. Figure 2.11 shows the individual processing steps applied to an visually aligned example image pair. A Sobel edge detector and a threshold of 0.3 times the maximum difference value were employed in this case. It can be observed that the edges extracted in the region of the enhancing lesion are quite different for the pre- and the post-contrast image, although the images are aligned. This demonstrates that edges are unsuitable features for guiding registration of contrast enhancing regions.

Figure 2.11: Example of extracting edges from a visually aligned DCE MR mammography image pair as suggested in [Lucht et al., 2000] produced by the author. The results of the processing steps are shown from left to right (a-d): a) original images, b) noise reduced images by applying a median filter to (a), c) extracted edges by applying a Sobel filter to (b), d) binary edges by thresholding the difference of the morphologically opened image (not shown) and the edge image (c). The bottom row shows the results of subtracting the first row from the second row, with gray intensities representing zero values. It can be observed that edges extracted from the pre- and the post-contrast images are very different within the enhancing region.
[Srikanchana et al., 2002] used the extracted edge at the breast skin for initial rigid alignment. Fibroglandular tissue was automatically segmented by fitting a Gaussian mixture model to the image histogram. The fibroglandular tissue was then partitioned into multiple objects. Local rigid alignment was based on the principle axes of these objects. This approach will suffer from fibroglandular tissue being misclassified as fatty tissue due to contrast enhancement.

### 2.3.2 Voxel Based Similarity Measures

This subsection introduces the definitions of voxel based measures for determining the similarity or the dissimilarity of two images. Optimization can be phrased as minimizing a cost function $C$ for all measures by negating the similarity measure $S$, i.e. $C = -S$.

Voxel based similarity measures assume a certain intensity relationship between the two registered images. Figure 2.12 illustrates the effects of motion and contrast enhancement on

![Figure 2.12: Effects of motion and contrast-enhancement on the intensity relation of DCE MR mammograms. The example was generated by the author. Column I shows an axial example slice of a visually aligned pre- and post-contrast image pair. Columns II-IV show the subtraction image and the joint histogram of the pre-contrast image and the following image: II) pre-contrast image rotated by $6^\circ$, III) post-contrast image rotated by $6^\circ$ and IV) post-contrast image. The histograms are presented using the negative logarithm of the joint probabilities, i.e. $-\log(p_T(a,b))$, in order to improve visibility of low probabilities.](image)
2.3 Image Matching Criteria

the intensity relationship of a pair of visually aligned DCE MR mammograms using joint
histograms\(^5\). The joint histogram of aligned, identical images will have zero entries every-
where apart from the diagonal (not shown). Noise or misalignment (column II) spreads
the distribution around the diagonal. Contrast enhancement causes a cluster of off-diagonal
entries (column IV). This cluster cannot be distinguished from off-diagonal entries caused
by misalignment (column III). This example illustrated the difficulty of defining a similarity
measure for contrast-enhanced images which penalizes image differences due to patient mo-
tion but not due to contrast change. Voxel based similarity measures which were employed
for DCE MR mammography registration are described next.

**Sum of Squared Differences (SSD)** is a very popular symmetric image dissimilar-
ity measure. The basic assumption is that both images are identical when registered, i.e.
\(B(T(x)) = A(x)\). The same assumption is made by the original optical flow registration
method [Horn and Schunck, 1981], i.e. changes in image intensities are only due to motion.
SSD is given by

\[
C_{SSD}(A, B^T) = \frac{1}{N^T} \sum_{x \in \Omega^T} [A(x) - B(T(x))]^2. \tag{2.30}
\]

SSD is the optimum measure if the images differ only by Gaussian noise [Viola, 1995] but is
sensitive to outliers. Intensity changes, due to the injection of contrast agent, clearly violate
this assumption of intensity preservation.

**Standard Deviation of Subtraction Image** is given by

\[
C_{STDS}(A, B^T) = \sqrt{Var(A - B^T)} \tag{2.31}
\]

where \(A - B^T\) is the subtraction image and \(Var()\) is the variance as defined in Notation.
The assumption made when using this symmetric dissimilarity measures is that the images
differ only by a constant term, i.e. \(B(T(x)) = A(x) + \alpha\). Again, this assumption is violated
due to regional contrast changes.

**Ratio Image Uniformity (RIU)** was proposed by [Woods et al., 1992] for the regis-
tration of images from different modalities. The assumption made when using RIU is that
the images differ by a factor, i.e. \(B(T(x)) = \beta A(x)\). RIU is defined as

\[
C_{RIU}(A, B^T) = \frac{\sqrt{Var(R^T)}}{R^T} \quad \text{where} \quad R^T(x) = \frac{B(T(x))}{A(x)} \quad \text{for} \quad x \in \Omega^T. \tag{2.32}
\]

\(^5\)The joint histogram of the images \(A\) and \(B^T\) is a 2D matrix whose \((a,b)\)th element records the probability
that \(A(x) = a\) and \(B(T(x)) = b\) for \(x \in \Omega^T\).
$C_{RIU}$ is an asymmetric dissimilarity measure and is independent of global intensity scaling. RIU is also referred to as variance of intensity ratios (VIR). $C_{RIU}$ is likely to be less well suited as a similarity measure for DCE MR mammography, since contrast change will increase the variance of the ratio image.

**Correlation Coefficient (CC)** measures the linear dependence between two datasets and hence is used when a linear relationship between the registered images is assumed, i.e. $B(T(x)) = \beta A(x) + \alpha$. CC is a symmetric similarity measure and is defined as

$$S_{CC}(A, B^T) = \frac{\text{Cov}(A, B^T)}{\sqrt{\text{Var}(A)\text{Var}(B^T)}},$$  \hspace{1cm} (2.33)

where $\text{Cov}(A, B^T)$ is the covariance of $A$ and $B^T$, and $\text{Var}(A)$ is the variance of $A$ as defined in Notation. $S_{CC}$ can take values from -1 (maximum negative correlation) to 1 (maximum positive correlation) and is either minimized or maximized depending on the expected relationship. CC is also called normalized cross correlation since it can be thought of as a normalization of the cross correlation measure $\sum_{x \in \Omega} A(x)B(T(x))$. Note that local contrast-enhancement does not provide a linear relationship between images.

**Correlation Ratio (CR)** was introduced for multimodal image registration by [Roche et al., 1998]. This measure assumes that the intensities of the images are related by some function $f$, i.e. $B(T(x)) = f(A(x))$. No assumptions are made about the nature of $f$. CR measures the functional dependency between two random variables. Let $\Omega_i$ be the iso-set of voxels in $A$ that have intensity value $i$. The CR is then defined as

$$S_{CR}(B^T | A) = 1 - \frac{\text{Var}(B^T - E(B^T | A))}{\text{Var}(B^T)} = 1 - \frac{\sum_i N_i \text{Var}_i(B^T)}{N^T \text{Var}(B^T)},$$  \hspace{1cm} (2.34)

where $E(B^T | A)$ denotes the conditional expectation of $B^T$ given $A$, $\text{Var}_i(B^T)$ is the variance of the voxels in $B^T$ that correspond to $\Omega_i$ and $N_i$ equals the number of voxels in $\Omega_i$. Note that $E(B^T | A)$ is the optimal function for minimizing $\text{Var}(B^T - f(A))$. CR is asymmetric, ranges from 0 (no functional dependence) to 1 (purely deterministic dependence). $S_{CR}(B^T | A)$ is independent to multiplicative changes in $B^T$. CR is closely related to the partitioned intensity uniformity (PIU) introduced by [Woods et al., 1993]. CR was not employed by any of the reviewed studies. It was, however, used in this work in Chapter 6 and is therefore included in this section for completeness.

**Joint Entropy (JE)** measures the amount of information provided by two combined images and is given by

$$C_{JE}(A, B^T) = H(A, B^T) = - \sum_a \sum_b p^T(a, b) \log p^T(a, b),$$  \hspace{1cm} (2.35)
where $p^T(a,b)$ is the joint probability that $A(x) = a$ and $B(T(x)) = b$ for $x \in \Omega^T$. $C_{JE}$ is minimized and equal to zero if $p(a,b) = 0$ for all but one intensity combination. $C_{JE}$ is maximized if all intensity combinations are equally likely. Both, [Collignon et al., 1995b] and [Studholme et al., 1995], proposed to use JE as an image dissimilarity measure. The main drawback of JE is that it is very sensitive to the overlap region $\Omega^T$. JE was not employed by any of the reviewed studies. It was, however, used in this work in Chapter 6 and is therefore mentioned.

**Mutual Information (MI)** measures the amount of information two images contain about each other. MI was independent and simultaneously pioneered by [Collignon et al., 1995a] and [Viola and Wells, 1995]. MI is a symmetric similarity measure and is defined as

$$S_{MI}(A,B^T) = H(A) + H(B^T) - H(A,B^T)$$

where $H(A) = -\sum_a p(a) \log p(a)$ is the marginal entropy of image $A$, with $p(a)$ being the probability of $A(x) = a$ for $x \in \Omega^T$. $H(A,B^T)$ is the joint entropy of $A$ and $B^T$ as defined in equation (2.35). $S_{MI}$ is minimized and zero if $A$ and $B^T$ are independent. $S_{MI}$ is never greater than the marginal entropies. By taking the marginal entropies into account, MI is less sensitive to the overlap region than JE. MI is very general since it assumes that a predictable relationship exists between the intensities of the two images.

**Normalized Mutual Information (NMI)** was proposed by [Studholme et al., 1999] to improve the robustness of MI to changes in the image overlap. NMI is given by

$$S_{NMI}(A,B^T) = \frac{H(A) + H(B^T)}{H(A,B^T)},$$

and was found to be considerably more robust for intermodality brain registrations when overlap changes are substantial [Studholme et al., 1999].

**Pharmacokinetic Model Fit** The residual error after fitting a pharmacokinetic model to the image sequence was used as an image matching criterion in [Hayton et al., 1997]. The model function for an instantaneous injection was expressed as

$$M(t) = \frac{A}{a-b} \left( \exp^{-bt} - \exp^{-at} \right)$$

where $A,a,b$ are the free parameters representing the amplitude, the exponential rise rate and the exponential fall rate respectively. The approximated distance of the data points to the model curve was used as residual measure. Model fitting was based on the sum of the squared residuals and penalty terms to prevent unrealistic parameters. Without significant contrast changes, residuals were dominated by noise, and convergence of the data fitting was problematic.
2.3.3 Summary

This section discussed the image matching criteria employed for DCE MR mammography registration. Many of these measures are inadequate for this task because they assume unaltered edges or penalize intensity changes. The intensity changes due to the injection of contrast agent are best reflected by pharmacokinetic models. However, model fitting within the unenhanced image regions was reported as problematic and the low temporal image resolution prevents the use of a sophisticated model. Adjusting the pre-contrast image intensity according to the tissue type may help if a segmentation is available. Of the remaining matching criteria, (normalized) mutual information makes the most general assumption and hence is likely to cope best with intensity changes.

Having decided on a transformation model and an image matching criterion, the remaining task is to develop a strategy for finding amongst all admissible transformations the one which maximizes the image matching criterion. This will be described next.

2.4 Optimization Strategies

This section describes the main characteristics of optimization methods employed for DCE MR mammography registration. Detailed descriptions of the discussed methods and additional optimization techniques can be found, for example, in [Press et al., 1992; Haykin, 1998].

The cost function is a $n$ dimensional function for a transformation model with $n$ degrees of freedom and is generally highly non-linear with many local minima. These minima can represent good local matches, can stem from interpolation artifacts, or can originate from changes in the region of image overlap. Filtering the images to reduce noise, multi-resolution approaches and normalizing the cost function with respect to the overlap region can reduce the number of local minima.

Yet, the global minimum might not necessarily provide the most accurate alignment either. A global minimum, for example, may be achieved by a transformation such that only air is within the region of image overlap. Reducing the search space to the maximum expected displacements may prevent such cases and reduces the search time. Success will depend on the true optimum being the global minimum in this reduced search space.

The size of the search space and the computationally expensive evaluation of the cost function may anyway prohibit methods based on exhaustive search or probabilistic exploration. Iterative methods are often used instead, where an initial estimate is improved by exploring the cost function in the local neighbourhood of the search space. The correct
transformation is found when the initial estimate is within the capture range\(^6\) of the true optimum. The extent of the capture range depends on the cost function, the image content, the field of view, and cannot be known \textit{a priori} [Studholme \textit{et al.}, 1997].

### 2.4.1 Direct Methods

The registration parameters are straightforwardly computed with direct methods. Examples include the least squares fitting of a transformation model to a set of displacements, as described in Section 2.2.1.2.

### 2.4.2 Discrete Methods

Discrete methods can be applied when the search space is sampled for a finite discrete set of transformations.

\textbf{Exhaustive Search} assesses all possible transformations and is only applicable to finite, small search spaces. It is guaranteed to find the best of these transformations.

\textbf{Genetic Algorithms} are based on iteratively assessing a random set of transformations, combining the parameters of the best and randomly perturbing them by a small amount. It is likely, but not guaranteed, that genetic algorithms converge to the global optimum.

### 2.4.3 Iterative Methods

The general strategy of iterative methods is to explore the search space in relatively small steps, where each step is guided by the local change of the cost function. This can be described by

\[
T_{k+1} = T_k - t_k = T_k - \alpha_k s_k,
\]

where \(t_k\) denotes the transformation update and \(T_0\) is the initial transformation. It is often useful to split the transformation update \(t_k\) into a step length \(\alpha_k\) and a search direction \(s_k\). Various methods arise depending on the chosen local change measure.

\textbf{Gradient Descent Algorithms} are very popular methods [Haykin, 1998]. The first partial derivative of the cost function \(C\) with respect to each parameter at \(T_k\), i.e. \(\nabla C(T_k)\), is used as search direction \(s_k\). Gradient descent methods based on line-search determine \(\alpha_k\) by minimizing \(C(T_k - \alpha_k \nabla C(T_k))\) with respect to \(\alpha_k\).

\textbf{Newton and Newton-Raphson Method} is based on minimizing the quadratic approximation of the cost function around the current transformation estimate \(T_k\) to speed up convergence. The search direction \(s_k\) is determined using the Hessian matrix \(\mathcal{H}C(T_k)\), denoting the second partial derivative of the cost function with respect to the transformation.

\(^6\)Region \(\mathcal{R}\) of the search space is the capture range of \(r\) if every initial estimate \(p \in \mathcal{R}\) converges to \(r\).
parameters at $T_k$, thus $s_k = [HC(T_k)]^{-1} \nabla C(T_k)$. In the Newton method, step length $\alpha_k$ is usually determined by a line search, while the Newton-Raphson method uses $\alpha_k = 1$.

**Gauss-Newton Method** is applicable to a cost functions which can be described by a sum of squared individual cost functions, i.e. $C(T) = \frac{1}{2} \sum c_i(T)^2$. The first order approximation of $c(T) = [c_1(T) \ c_2(T) ... c_n(T)]^T$, when $T$ is decreased by $t = [t_1 \ t_2 ... t_n]$, is given by $c(T - t) = c(T) - \nabla c(T)t$, where the Jacobian matrix $\nabla c(T)$ has elements $\nabla c_{ij}(T) = \frac{\partial c_i(T)}{\partial T_j}$. The cost function $C(T - t)$ is minimized if all individual $c_i$ are minimized, i.e. $c(T - t) = 0$, which implies that $\nabla c(T)t = c(T)$. The best approximation of $t$ in the least squares sense at the $k$th iteration is then given by $t_k = \left(\nabla c(T_k)^T \nabla c(T_k)\right)^{-1} \nabla c(T_k)^T c(T_k)$.

**Conjugate Gradient Descent Methods** belong to the class of second-order optimization methods. The aim is to speed up gradient descent while avoiding the computational requirements for inverting the Hessian matrix in the Newton method. The main idea is to follow gradient directions which do not interfere with each other for the given problem. That means, for minimizing the quadratic function $f(x) = \frac{1}{2}x^T Ax - b^T x + c$, when $A$ is a symmetric and positive definite matrix, the search directions are given by $s_k$ such that $s_k^T As_k = 0$ for all $j$ and $k$ when $j \neq k$. The explicit knowledge of $A$, i.e. the Hessian matrix, can be avoided by sequentially determining the search directions using $s_k = -\nabla C(T_k) + \beta_k s_{k-1}$ and $\beta_k = \frac{\nabla C(T_k)^T \nabla C(T_k)}{\nabla C(T_{k-1})^T \nabla C(T_{k-1})}$ [Haykin, 1998].

### 2.4.4 Multiresolution

Multiresolution methods start by registering low resolution images using transformation models with low number of degrees of freedom. This initial transformation is then iteratively refined for images with improved resolution and transformation models with increased number of parameters. The aim is to avoid local minima and to reduce the search space of the finest, most expensive level.

### 2.4.5 Summary

This section described the optimization strategies employed for DCE MR mammography registration. Non-rigid models with a large number of transformation parameters require in general iterative optimization methods. Robustness and efficiency can be improved by multiresolution approaches.

The last three sections described the main components of registration algorithms, namely the transformation model, the image matching criterion and the optimization strategy. Having developed a registration method, the final task is to assess its performance, which will be discussed next.
### 2.5 Evaluation Schemes

Without careful validation, a registration method cannot be accepted as a clinical tool. Validation criteria include accuracy, reliability, robustness, fault detection, functional complexity and clinical use [Maintz and Viergever, 1998; Jannin et al., 2002]. An objective assessment of the registration outcome is, however, a difficult task because the true transformation (termed here the ground truth) is generally unknown. In this section, schemes are reviewed which were used to evaluate DCE MR mammography registration.

#### 2.5.1 Target Registration Error

The accuracy of the registration can be quantified if the ground truth transformation $T_g$ is known. The registration error of $T$ at point $x$ is then given by

$$TRE(x) = ||T(x) - T_g(x)||,$$

where $TRE$ stands for target registration error to emphasize that the error is measured at anatomical positions of interest [Fitzpatrick, 2001]. Apart from the rare case of simple translational errors, TRE varies spatially across the image. Error images, showing this spatial TRE distribution, can be very helpful in investigating individual results. The registration accuracy for a region of interest can be summarized by the mean, median, RMS, 95th percentile or maximum of the regional TRE distribution. Unfortunately, in most cases the ground truth is unavailable and TRE cannot be calculated.

#### 2.5.2 Gold Standards

A registration method whose accuracy is known to be high can be used to provide a gold standard transformation [Fitzpatrick, 2001]. The results of other registration methods can then be compared against the gold standard result instead of the ground truth.

Rigid registration based on the location of bone-implanted markers is for example a gold standard for the registration of intermodality head images for surgical patients. Its TRE at locations of interest is about 0.5mm [Maurer et al., 1997]. After removal of the markers from the images, registration strategies which do not rely on these markers can then be tested against this gold standard [West et al., 1997]. This approach has been shown to be effective for rigid intermodality registration, while a much higher accuracy of the gold standard is required for intramodality registration or non-rigid registration.

Gold standard image pairs can also be created by computer simulations. Both images could be produced by modelling the anatomy and the imaging process. More common is however that a known geometric transformation is applied to an image to generate a misaligned
second image. While this approach provides an accurate gold standard transformation, the resulting image pairs are often not very realistic. In the case of DCE MR mammography, simulations may lack the imaging artifacts caused by motion, may not model the intensity changes due to the injection of contrast agent or may have unrealistic deformations.

2.5.3 Visual Inspection

The most obvious method for assessing the registration outcome is through visual inspection. The assessment may be facilitated by interactive displays of colour overlays, subtraction images, interleaved images, surface renderings or continuous film sequences (cine mode). Visual inspection is observer dependent, labour intensive, non-reproducible and generally qualitative. Explicit manual identification of corresponding anatomical features, like landmarks or regions, enables quantitative assessment (e.g. distance of corresponding landmarks or overlap of segmented regions). This demands a greater effort than qualitative assessment for a single inspection, but provides a basis for automatically quantifying many different registration methods. Corresponding structures need to be established with a sufficient accuracy to make this approach useful.

The ability of visually detecting or identifying aligned structures depends greatly on the image content. Difficulties will increase as images differ due to noise, inhomogeneities, distortions, small rigid motion, large non-rigid motion, the injection of contrast agent, anatomical changes, the type of modality and the imaged subject. For DCE MR mammograms, the injection of contrast agent prevents identification of corresponding features in the contrast-enhanced regions.

Visual inspection can serve as a gold standard as long as its accuracy is known to be high. A few studies have assessed the sensitivity of the radiologist to detect or estimate misalignments. [Wong et al., 1997] misaligned registered MR and PET images by 8 translations (-1mm, -2mm, ..., 4mm) or by 8 rotations (-1°, -2°, ..., 4°) in each direction. The observers always detected misalignments greater than 2° and 2mm in the image plane; and 3°-4° and 3mm in the slice direction. [Fitzpatrick et al., 1998] assessed the ability of radiologists to estimate the error of rigid body misalignments for MR and CT head images relative to 6 thresholds (1-6mm). The voxel dimensions were approximately 0.4×0.4×3.0mm³ (CT) and 1×1×1.6mm³ (MRI). They found that errors of 2mm or more were well estimated. In [Denton et al., 2000], radiologists were able to detect translational misalignments greater than 0.2mm for sequential MR brain images with voxel dimensions 1×1×1.8mm³ when tested on 10 translations (0.05-0.50mm). These studies show that the ability of radiologists to detect
or estimate misalignments strongly depends on the image content.

### 2.5.4 Consistency Registration Error

Without the availability of a realistic gold standard, registration algorithms are frequently assessed by measuring the consistency of the transformations from individual registrations [Freeborough et al., 1996; Woods et al., 1998; Holden et al., 2000]. For three images $A$, $B$ and $C$ of a subject, a point $x$ in $A$ can be followed from $A$ to $B$ to $C$ and back to $A$ by individually registering $A$ to $C$, $C$ to $B$ and $B$ to $A$. The consistency registration error (CRE) of this so-called registration circuit is then given by

\[
CRE(x) = ||T_{AC}(T_{CB}(T_{BA}(x))) - x||,
\]

where transformation $T_{BA}(x)$ maps point $x$ in image $A$ to point $T_{BA}(x)$ in image $B$. If the errors of the individual registrations are uncorrelated, then the error of the single registrations for a circuit of $n$ images is given by $CRE/n$. This assumption is however violated by the fact that successive registrations share an image and that the same registration method was applied. Dividing CRE by $n$ will therefore generally underestimate the individual registration errors. For example, the CRE from the forward and backward rigid registration of a pair of DCE MR mammograms could be low although the individual registration errors may be high due to non-rigid deformations.

### 2.5.5 Image Matching Criteria

By their nature, voxel based image similarity measures cannot quantify spatial registration accuracy. Their use as surrogate measures is questionable, since they are often not independent from the employed optimization criterion, and their assumed positive relationship with registration accuracy may (partially) be violated (e.g. shrinking an enhancing lesion decreases the SSD but increases TRE).

The registration accuracy could be assessed by extracting features and pairing them. But inaccuracies of feature extraction or the use of these features during the registration will put limits to this approach.

### 2.5.6 Clinical Usefulness

The ultimate goal of DCE MR mammography registration is to improve the accuracy of breast cancer detection and diagnosis. Depending on the severity of the misalignment, registration may offer several types of improvements. When severe misalignments hinder the film reading, registration may save the additional repeat scan. The reading time gets faster because of better quality subtraction images. The false positive rate can be reduced by
avoiding misinterpretation of motion artifacts as enhancements. Discrimination of benign and malignant lesions may improve due to higher authenticity of the lesions’ features.

2.5.7 Summary

This section reviewed strategies for assessing DCE MR mammography registration results. Estimation of the registration accuracy is problematic, because the true transformation is unavailable. Image intensity changes make visual inspection unreliable in the region of enhancing structures. The deformable nature of breasts and the changes of the image intensities make the creation of gold standard transformation challenging. Consequently, evaluation of DCE MR registration results is more complicated than it may seem.

All concepts required to review DCE MR mammography registration methods have now been introduced. The individual methods are reviewed in the next section.

2.6 Review of Registration Methods

This section reviews the individual DCE MR mammography registration methods in chronological order. Readers familiar with the topic may skip the individual reviews and read instead the summary provided in Section 2.6.2.

2.6.1 Individual Methods

- The aim in [Davis et al., 1995; Davis et al., 1997] was not primarily to register DCE MR mammograms, but to find a realistic transformation model for a given set of corresponding anatomical landmarks. For this purpose, they introduced the elastic body spline (EBS) as defined by (2.12). The performance of the EBS was compared with that of an affine model (2.1), a thin plate spline (2.9) and a 3D volume spline (Table 2.1). Tests included firstly 4 plain7 MR breast images with simulated deformations (a constant volume, 3D 2nd-order Taylor series warp) and landmarks selected at random locations. The second test set included 3 plain MR breast images with actual deformations. Pairs of control points were automatically determined using a block matching method based on the local correlation coefficient (2.33) and an exhaustive search for the best local translation. The performance was assessed on the basis of 3 image matching criteria, namely: mean squared error, mean absolute error and correlation coefficient. The EBS achieved better results than the other interpolation methods in all but one case. The difference was, however, small for images with actual deformations and decreased with the number of landmarks.

---

7Plain MR images refers to not contrast-enhanced MR images.
• [NessAiver et al., 1995; NessAiver et al., 1996] registered rigidly (2.2) the individual slices of 3D DCE MR mammograms such that the standard deviation of the subtraction image (2.31) was minimized. Optimization was based on gradient descent. This improved the employed image dissimilarity measure and the visual appearance of the subtraction image in comparison to the 3D rigid registration in [NessAiver et al., 1995]. This method does not account for any out-of-plane motion or in-plane deformation and will create discontinuous transformations between slices.

• [Kumar et al., 1995; Kumar et al., 1996] employed a 3D affine transformation (2.1) for global alignment followed by 3D local flow estimation. Images were filtered by a Gaussian, a Laplacian and an oriented Gaussian filter. The sum of squared differences (2.30) of these filtered data sets was employed as a dissimilarity measure. A flow vector was calculated for each voxel using a Gauss-Newton optimization method and a multi-resolution approach. The final transformation was smoothed by assuming that the image velocities are constant within a region of 5 x 5 x 5 voxels. Evaluation consisted of visual inspection of 6 patient cases using subtracted images and maximum intensity projections. A reduction of motion artifacts can be observed in the examples presented in the paper. The method assumes, however, that edges are preserved, which does not hold for enhancing regions.

• [Zuo et al., 1996] registered breast images using a 3D rigid transformation (2.2) and by minimizing the variance of the ratio image of two images (2.32). Parameters were iteratively determined using a modified Newton-Raphson method. When tested on 5 patients, registration reduced visually motion artifacts (subtraction images and cine mode), reduced the variance across the breast boundary of the subtraction images and improved the signal-to-noise-ratio (SNR) within the lesion. Simulation of an enhancing lesion without an induced transformation produced rotation angles of less than 0.05° and translations of less than 0.02mm. This method does not account for any non-rigid transformation.

• [Undrill et al., 1996] registered 2 patients by interpolating manually identified landmarks by 3D orthogonal polynomials (see Section 2.2.1.2). Results were compared with two automatic methods which both employed a 3D rigid transformation model (2.2) plus scaling. The first was based on iteratively computing the principle axes for each image using the second order invariant moments ([Gonzalez and Woods, 1993]). The second method minimized the variance of the ratio image (2.32) using conjugate gradi-
ent descent. Based on visual inspection, interpolation by polynomials was most effective in improving the quality of the subtraction images. This result supports the claim that a rigid transformation is not sufficient for aligning breast images. Assessment was however only based on 2 patients.

- [Huwer et al., 1996] employed a modified Kohonen network model for registering MR breast images. The network’s architecture consisted of one neuron per image pixel. The weights of each neuron stored the displacement vector information. Winner neurons were determined according to the best matching image features (a weighted sum of various image features as defined in (2.29)) for all pixels in the reference image which were not in a grey-level plateau\(^8\). The learning rule was modified such that neighbouring neurons of a winner were collinearly updated according to the winner’s displacement vector. This update was weighted by a Gaussian function that defined the neighbourhood of the winner. The radius of this Gaussian function was decreased as learning progressed. It was reported that based on visual inspection of the subtraction images, the approach produced reliable results for 14 patients, even for relatively strong dislocations. The images features assume preservation of image intensity and image edges, which does not hold in regions of enhancement.

- [Davey et al., 1997] observed for 9 patients that a 3D rigid registration (2.2) per breast provided visually better subtraction images than if the registered image contained both breasts. Optimal rigid transformation parameters were determined using a genetic algorithm. This result supports the perception that both breasts move independently.

- [Hayton et al., 1997] employed a 2D optical flow inspired approach which minimized the error when fitting a pharmacokinetic model (2.38) to the image sequence within the segmented breast. The transformation was regularized by averaging the displacement field locally during gradient descent. The algorithm achieved for one case a 2D mean registration error of about 0.7mm for images misregistered by global translations of 1.6mm and 2.2mm. Convergence of the model fitting depended on significant enhancement, since the residuals were otherwise dominated by noise. Assessment was based on only one example.

- The aim in [Hess et al., 1998] was to develop a fast method suitable to reduce the in-plane motion artifacts of 6mm thick axial DCE MR mammogram slices. For a subset

---

\(^8\)A plateau is defined as 3x3 region with a local brightness gradient which was below a certain threshold.
of voxels, 2D image blocks were matched using the sum of squared differences (2.30) and exhaustive exploration of all local translations. A continuous transformation was derived by Voronoi based interpolation. Low resolution subtraction images were shown for only one patient and no computation time was given. This approach is likely to produce an inaccurate alignment in regions of contrast change because of the employed optimization criterion.

- Rueckert et al., 1998; Rueckert et al., 1999a; Rueckert et al., 1999b] registered DCE MR mammograms using an initial 3D rigid transformation model (2.2) followed by a free-form deformation (FFD) model. The FFD model is based on a regular grid of control points approximated by cubic B-splines (2.13). The transformation parameters (translations, rotations and control point displacements) were determined on the basis of maximizing normalized mutual information (2.37) using a multi-resolution gradient descent approach. The 3D equivalent of the thin plate spline bending energy (2.22) was employed as a regularizer to keep the transformation smooth. The quality of the subtraction images before and after rigid, affine and non-rigid FFD registration was visually ranked for 54 DCE MR mammograms by two radiologists [Denton et al., 1999]. This study showed that the non-rigid FFD registration provided statistically significant (at the 1% level) better-quality subtraction images. The influence of interpolation and registration methods on the diagnostic accuracy of enhancing MR breast lesions was assessed in [Sonoda, 2003]. Here, classification was based on radiologists’ reading or on stepwise linear discriminant analysis of features extracted from a manual segmentation of the lesion. Diagnostic accuracy was assessed by the area under the receiver operating characteristics curve (AUC). For a set of 40 enhancing lesions and classification based on 3D features, results improved statistically significantly at the 5% level after rigid, affine or 10mm non-rigid FFD registration using sinc interpolation in comparison to unregistered images. Rigid and affine registration provided with 0.81 a markedly better AUC value than 10mm non-rigid FFD registration (AUC=0.73). This is the only study which assessed the influence of registration on the classification of MR breast lesions. The registration method itself has the advantage of using computationally less expensive interpolation functions (B-splines), and an optimization criterion which is likely to be less influenced by intensity changes than others.

- Meyer et al., 1998] optimized the position of manually selected corresponding control points of a 3D thin plate spline (2.9) according to global mutual information (2.36).
Gradient descent from different starting estimates was employed for optimization. The aim of this study was to find the minimal set of control points. Retrospectively, a negative correlation between curvature of the cost function and variance in the final control point position was found after several runs with randomly displaced initial positions. The local gradient magnitude of the cost function at initial positions was however uncorrelated with the variance in the final control point position. Subtraction images were shown for one patient. Drawbacks of this methods are the manual placement of control points and the global interpolation function.

- [Fischer et al., 1998; Fischer et al., 1999a] employed a regular, trilinearly interpolated grid of control points (with usually 16 voxels spacing) as their transformation model. Control point positions were iteratively adjusted (by gradient descent) such that the correlation (2.33) between the affected 8 subvolumes of the source and reference image were maximized. Based on the visual appearance of the subtraction images, the registration performed well for 1 volunteer and 3 patients, and failed for 1 patient with an initial motion of more than 10mm. The employed optimization criterion assumes a linear relationship which may be violated by contrast changes. Assessment was based on only 3 patients.

- [Lucht et al., 1998; Lucht et al., 2000] first rigidly aligned the slices of DCE MR mammo-grams such that the overlap of the thresholded images was maximized. Local gray-level changes (edges) were then extracted as described in Section 2.3.1. The displacement vectors of the features were optimized such that the area of overlapping edges was maximized. Bi-linear interpolation was employed to derive 2D displacement vectors for regions without features. Finally, the transformation was smoothed by 3D Gaussian filtering. The results were visually assessed by tracing anatomical landmarks in a continuous film sequence for 20 patients. The visual alignment after non-rigid registration was found to be better than without registration. The registration accuracy was estimated to be about 1 pixel for 19 patients. The non-rigid registration achieved for 8 patients better and for 2 patients equally good results than rigid registration when ranking the visual appearance of 10 image pairs [Brückner et al., 2000]. The method assumes preservation of edges which will not hold in the region of contrast enhancement.

- The aim of [Krishnan et al., 1999] was to correct for patient motion in k-space prior

\[^9\text{MR data is acquired using frequency and phase encoding and k-space contains this data. MR images are}\]
to offline keyhole reconstruction. A spatial translation in image-space causes a phase shift in k-space. Linear phase shifts were estimated by determining the average phase deviation for each spatial frequency and fitting a representative slope. Phase wrapping as well as constant phase shifts were taken into account. Robustness to contrast enhancement was shown by imaging a DCE breast phantom in two different positions. For 64 symptomatic patients, the mean subtraction edge artifacts\textsuperscript{10} were never increased and for most cases decreased by motion correction. This method compensates only for rigid motion.

- \textit{[Hayton et al., 1999]} employed mutual information (MI) (2.36) to measure at equally spaced pixels the local image similarity for a set of possible 2D displacements. The MI measure of a displacement was then multiplied by the maximum MI of similar displacements achieved in its surrounding, weighted by a spatial Gaussian function. After repeating this proceeding from a coarse to a fine spacing, the displacement with the maximum MI was selected per sampled pixel. A 2D cubic B-spline grid (2.13) was then fitted to these displacements such that the residual errors, weighted against the bending energy (2.8) of the splines, were minimized. This motion estimation process was then repeated on a finer scale. Results (deformed images, subtraction images and displacement fields) were shown for 4 breasts. For the patient with the most complex motion, the consistency error between the forward and backward registration result was at most 4 pixels (4mm) and in relevant areas 2 pixels (2mm). This method does not account for out-of-plane motion and its assessment was based on only 4 cases.

- \textit{[Vieira and Undrill, 1999]} compared the performance of two registration methods known as the automatic image registration (AIR) \textit{[Woods et al., 1998]} and the spatial parametric mapping (SPM) \textit{[Friston et al., 1995]}. In AIR, the parameters of 2nd- to 5th-order 3D polynomials (2.5) are optimized (using the Newton method) such that the variance of the ratio image (2.32) is minimized. In SPM, the parameters of a linear combination of 3D discrete cosine transform basis functions (2.6) are optimized (using the Gauss-Newton strategy) such that the sum of squared differences (2.30) weighted against the membrane energy (2.18) is minimized. Performance was assessed for a simulated misaligned image pair and for one patient. The misalignment was created by locally

\textsuperscript{10}The mean subtraction edge artifacts are defined by the mean intensity of a manually placed region which included all the edge artifacts of a representative mid-breast slice from the subtraction image.
shearing each breast parallel to the rib cage. Intensity changes due to the administration of contrast agent were not included. AIR employing 5th-order polynomials was superior in reducing the intensity changes of these misaligned images. Registration errors were not calculated. The subtraction images for one volunteer and one patient were visually most improved by AIR. Both methods employ image matching criteria which will be affected by intensity changes. Evaluation was limited due to the simple gold standard deformation, the lack of incorporating intensity changes, the omission of an accuracy analysis and the small dataset.

- [Wirth, 1999] employed 2D multiquadratic radial basis functions (2.7)\textsuperscript{11} to interpolate the 2D displacement of corresponding landmarks. Interior landmarks were selected manually, while landmarks on the breast to air boundary were extracted automatically from an automatic segmentation. Three MR slices for 1 patient were registered using 30 interior and 66 skin landmarks per 2D slice. Image similarity measures (SSD (2.30), CC (2.33)) were most improved by the multiquadratic functions in comparison to an affine (2.1) or a thin plate spline (2.9) transformation model. This manual method does not account for out-of-plane motion. The small number of cases and the use of image similarity measures as an evaluation criterion for contrast-enhanced images does not allow to draw general conclusions from the results.

- [Reichenbach et al., 2000; Reichenbach et al., 2002] employed mutual information (2.36) as an image similarity measure. 3D global translation parameters were first optimized by local exhaustive search using the 5 central slices of each individual breast. Equidistant pre-selected 2D slices were then individually registered to the 3D reference image using 2 out-of-plane rotation angles (no rotation in the axial plane). The rotation angles for the remaining slices were determined by linear interpolation or by extrapolation using a linear fit to the rotation angles of the pre-selected slices. The performance of the method was assessed using the image matching criterion (mutual information). Registrations were performed for plain MR images taken before and after repositioning either a rigid phantom or 3 volunteers. The employed transformation model does not allow for any in-plane deformation. Evaluation was based on the employed optimization criterion, plain MR images and a small dataset, and hence is insufficient.

- [Rohlfing and Maurer, 2001; Rohlfing et al., 2003] altered the registration method

\textsuperscript{11}The locality parameter \(c\) was set to the mean radial distance between corresponding control points.
from [Rueckert et al., 1999b] by regularizing the transformation to preserve volume (2.21) instead of minimizing the 3D bending energy (2.22). Maximum intensity projections of the subtraction images of 17 patients were visually assessed for motion artifacts. For each patient, regularization was balanced against image similarity NMI (2.37) by choosing the smallest weight which kept lesion volume change within 2%. The unconstrained non-rigid registration results were ranked best for all 17 cases, while the rigid registration results were ranked worst for 16 cases. Regularization of the non-rigid registration by the volume-preservation constraint achieved a slightly better ranking than regularization by the 3D bending energy (9 cases better, 4 cases equal, 4 cases worse). Visual appearance of the worst volume-preserving case improved, while maintaining a 1% volume change, by alternating between an unconstrained and a volume-preserving constrained cost function during the non-rigid registration. This study confirms that implausible volume changes of the enhancing lesion are not detected by visual inspection.

- [Fischer and Modersitzki, 2002b; Fischer and Modersitzki, 2002a; Fischer and Modersitzki, 2004] registered a pair of DCE MR mammograms using sum of squared differences (2.30) weighted against curvature regularization of the displacement field (2.23). Optimization was based on gradient descent employing a finite difference approximation. The employed image similarity measure will penalize intensity changes. Evaluation was based on visual inspection of a single case and hence insufficient.

- [Srikanchana et al., 2002] firstly rigidly (2.2) registered a pair of DCE MR mammograms based on the extracted edge of the breast skin. Fibroglandular tissue objects were then extracted as described in Section 2.3.1. Subsequently, local rigid transformations were determined based on the principle axes of the objects. The individual transformations were then combined to a mixture of piece-wise rigid registrations. Thereafter, high-curvature points were extracted from the boundaries of the Fibroglandular tissue using corner detection. Finally, a three-layer multi-perceptron neural network was trained with the coordinates of these corresponding points to establish the parameters of a polynomial transformation (2.5). This method will suffer from fibroglandular tissue being misclassified as fatty tissue due to contrast change. Evaluation was based on visual inspection of a single case and hence inadequate.

- [Haber and Modersitzki, 2004b; Haber and Modersitzki, 2004a] based registration on the sum of squared differences similarity measure (2.30), an elastic regularization
term (2.27) and a constraint to keep local volume changes minimal (2.20). Optimization was based on discretization of the resulting partial differential equations and solving it using a Newton-type optimization method. The regularization weight was determined by visual inspection. Subtraction images were shown for one breast. Intensity changes are penalized by the employed similarity measure. Evaluation was again limited.

- [Xiaohua et al., 2005] proposed an iterative segmentation and registration approach. For a given transformation, the breast tissue was segmented into fatty, normal, benign and malignant tissue as described on Page 129. For a given segmentation, the images were then registered using a similarity measure that is proportional to the sum of squared differences (2.30) with the pre-contrast image adjusted for each class. No information was provided regarding the exact adjustment used. The transformation was regularized by penalizing changes of the transformation parameters in the local neighbourhood. This iterative process of segmentation and registration was then repeated until convergence. Optimization was based on iterative gradient descent. Results for one patient showed that the residual norm value of fitting a pharmacokinetic model was reduced after registration, an improved segmentation was reported and no problems with volume changes.

### 2.6.2 Summary

This section reviewed the published methods for registering DCE MR mammograms acquired during one visit. Their components and parameters are listed in Table 2.4 and 2.5.

Most methods employed 3D non-rigid transformation models to account for the deformability of breasts. The transformation models were rarely founded on relevant physical principles. Exceptions were linear elastic models and volume preservation constraints, which both take care of cross-effects from the individual components.

The administration of contrast agent presents a major difficulty for defining an adequate image matching criterion, since neither edges nor intensities are preserved in enhancing regions. Fitting a pharmacokinetic model could be the answer for regions with contrast change, but is problematic for unenhancing regions. Adjustment of the pre-contrast image intensity according to the tissue type may help if a segmentation is available. Mutual information and normalized mutual information are based on the least restrictive assumption and hence are likely to cope best of all other measures. None of the reviewed studies compared the performance of different matching criteria, because it is difficult to assess or quantify registration accuracy.
<table>
<thead>
<tr>
<th>Reference</th>
<th>ty(^a)</th>
<th>Transformation Model(^b)</th>
<th>Image Matching Criterion(^c)</th>
<th>Evaluation</th>
<th>#(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Davis et al., 1997]</td>
<td>J95</td>
<td>3D affine, 3D EBS</td>
<td>CC(image blocks)</td>
<td>induced warp, no CE, SSD, CC</td>
<td>7</td>
</tr>
<tr>
<td>[NessAiver et al., 1996]</td>
<td>A95</td>
<td>2D rigid</td>
<td>STD(subtraction image)</td>
<td>visual</td>
<td>1</td>
</tr>
<tr>
<td>[Kumar et al., 1996]</td>
<td>C95</td>
<td>3D affine, 3D local averaging</td>
<td>SSD(filtered images)</td>
<td>visual</td>
<td>12</td>
</tr>
<tr>
<td>[Zuo et al., 1996]</td>
<td>J96</td>
<td>3D rigid</td>
<td>RIU</td>
<td>visual, SNR(lesion), STD(breast border)(^2)</td>
<td>10</td>
</tr>
<tr>
<td>[Undrill et al., 1996]</td>
<td>C96</td>
<td>3D rigid + scaling</td>
<td>RIU</td>
<td>visual</td>
<td>3</td>
</tr>
<tr>
<td>[Huwer et al., 1996]</td>
<td>J96</td>
<td>2D Gaussian smoothing</td>
<td></td>
<td>visual</td>
<td>28</td>
</tr>
<tr>
<td>[Davey et al., 1997]</td>
<td>A97</td>
<td>3D rigid</td>
<td>?</td>
<td>visual</td>
<td>18</td>
</tr>
<tr>
<td>[Hayton et al., 1997]</td>
<td>J97</td>
<td>2D local averaging</td>
<td>pharmacokinetik model</td>
<td>visual, induced translation</td>
<td>1</td>
</tr>
<tr>
<td>[Hess et al., 1998]</td>
<td>C98</td>
<td>2D Voronoi interpolation</td>
<td>(\sqrt{SSD(image \ blocks)})</td>
<td>visual</td>
<td>2</td>
</tr>
<tr>
<td>[Rueckert et al., 1999b]</td>
<td>J98</td>
<td>3D rigid, 3D cubic B-spline grid, BE</td>
<td>NMI</td>
<td>visual ranking, diagnostic accuracy (40)</td>
<td>54</td>
</tr>
<tr>
<td>[Fischer et al., 1999a]</td>
<td>J98</td>
<td>3D linearly interpolated grid</td>
<td>CC(image blocks)</td>
<td>visual</td>
<td>8</td>
</tr>
<tr>
<td>[Lucht et al., 2000]</td>
<td>J98</td>
<td>2D rigid, 2D linear ip., 3D Gaussian sm.</td>
<td>edges</td>
<td>visual ranking (10 patients)</td>
<td>40</td>
</tr>
<tr>
<td>[Krishnan et al., 1999]</td>
<td>J99</td>
<td>3D translation</td>
<td>k-space phase deviation</td>
<td>phantom, mean edge artifacts</td>
<td>64</td>
</tr>
<tr>
<td>[Hayton et al., 1999]</td>
<td>J99</td>
<td>2D cubic B-spline grid, BE</td>
<td>MI(image blocks)</td>
<td>visual, consistency error</td>
<td>4</td>
</tr>
<tr>
<td>[Vieira and Undrill, 1999]</td>
<td>C99</td>
<td>3D polynomials; 3D DCT, ME</td>
<td>RIU; SSD</td>
<td>induced local shear, intensity ratio, visual</td>
<td>3</td>
</tr>
<tr>
<td>[Reichenbach et al., 2002]</td>
<td>J00</td>
<td>3D translation, 3D slice rotation</td>
<td>MI(2D slice)</td>
<td>phantom, MI, visual, no CE</td>
<td>3</td>
</tr>
<tr>
<td>[Rohling et al., 2003]</td>
<td>J01</td>
<td>3D rigid, 3D cubic B-spline grid, VP</td>
<td>NMI</td>
<td>visual ranking</td>
<td>17</td>
</tr>
<tr>
<td>[Fischer and Modersitzki, 2002a]</td>
<td>J02</td>
<td>curvature regularization</td>
<td>SSD</td>
<td>visual</td>
<td>1</td>
</tr>
<tr>
<td>[Srikanthana et al., 2002]</td>
<td>C02</td>
<td>rigid, polynomials</td>
<td>fibroglandular objects, corners</td>
<td>visual</td>
<td>1</td>
</tr>
<tr>
<td>[Haber and Modersitzki, 2004a]</td>
<td>J04</td>
<td>elastic regularization, VP</td>
<td>SSD</td>
<td>visual</td>
<td>1</td>
</tr>
<tr>
<td>[Xiaohua et al., 2005]</td>
<td>C05</td>
<td>local smoothness regularization</td>
<td>SSD, pre-contrast adjusted</td>
<td>visual, pharmacok. model fitting error</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 2.4: Overview of automatic registration algorithms for DCE MR mammography.

\(^a\)ty: abstract (A), conference (C), journal (J); y: last two digits of year of first publication

\(^b\)BE: bending energy, DCT: discrete cosine transform, EBS: elastic body spline, ME: membrane energy, VP: volume preservation, ip.: interpolation, sm.: smoothing

\(^c\)CC: correlation coefficient, (N)MI: (normalized) mutual information, RIU: ratio image uniformity, SSD: sum of squared differences, STD: standard deviation

\(^d\)\#number of assessed breasts
2.7 Proposed Validation Method

More than 20 methods have been proposed for the intra-visit registration of DCE MR mammograms in the last 10 years. Assessments of these methods relied mostly on qualitative visual inspection. Yet, the accuracy of observers to assess misalignments of DCE MR mammograms, especially within enhancing regions, is unknown. The visual appearance of registration results improved for transformation models with more degrees of freedom [Denton et al., 1999; Brückner et al., 2000; Rohlffing et al., 2003]. As the degrees of freedom increased, however, volume changes within the region of the enhancing lesion rose [Tanner et al., 2000;...]

---

Table 2.5: Overview of non-automatic (manual) registration algorithms for DCE MR mammography.

<table>
<thead>
<tr>
<th>Reference</th>
<th>ty&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Transformation Model</th>
<th>Image Matching Criterion</th>
<th>Evaluation&lt;sup&gt;b&lt;/sup&gt;</th>
<th>#&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Undrill et al., 1996]</td>
<td>C96</td>
<td>3D orthogonal polynomials</td>
<td>manual landmarks</td>
<td>visual</td>
<td>3</td>
</tr>
<tr>
<td>[Meyer et al., 1998]</td>
<td>C98</td>
<td>3D thin plate spline</td>
<td>manual landmarks, MI</td>
<td>visual, MI</td>
<td>1</td>
</tr>
<tr>
<td>[Wirth, 1999]</td>
<td>T99</td>
<td>2D multi-quadratic</td>
<td>skin contour landmarks,</td>
<td>visual, CC, SSD</td>
<td>1</td>
</tr>
</tbody>
</table>

<sup>a</sup>t: conference (C), thesis (T); y: last two digits of year of first publication

<sup>b</sup>CC: correlation coefficient, MI: mutual information, SSD: sum of squared differences

<sup>c</sup>number of assessed breasts

Evaluation was based on image similarity, visual inspection or simulated deformations. Image similarity is a surrogate measure often used for optimization and cannot provide information about registration accuracy. Visual inspection ranged from showing a single example to letting radiologists rank the quality of subtraction images for a relative large dataset. Unfortunately, it was found that visual inspection does not identify implausible volume changes of the enhancing lesions. Its usefulness is therefore limited. Three studies employed simulated misregistered image pairs as a basis for performance assessment. These simulations were, however, rather unrealistic due to oversimplistic deformations and a lack of accounting for intensity change. Only one study evaluated the clinical usefulness of registration.

Many registration methods have been proposed for DCE MR mammography. Comparison of these methods is however hindered by the lack of an adequate evaluation scheme. To improve this situation, a novel validation method was developed in the context of this thesis. This validation method is put forward next.
Rohlfing et al., 2003]. The implausible amount of the visually undetected volume changes, between [-19,18]% and [-78,-1]% respectively, suggests that observers have problems to identify misregistrations within the enhancing regions.

Quantitative evaluations were restricted to image matching criteria, consistency measures and to registration errors from simulated transformations. Assessments based on image matching criteria do not provide any information about the registration accuracy. Conversely, consistency measures can provide a lower limit to the registration error, but will fail to detect any bias of the method. The transformations simulated to date were simple or unrealistic, and no contrast change was modelled.

No ground truth exists for the intra-visit registration of DCE MR mammograms. The approximation of a gold standard transformation from fiducial markers\(^\text{12}\) requires that these are close to the region of interest, since breasts deform locally. Skin markers are restricted by the amount skin slides over the underlying tissue, while markers implanted into soft tissue are generally instable. Natural anatomical landmarks in non-enhancing regions may serve as fiducials. Accuracy of such a gold standard will depend on the error to localize the landmarks and on the accuracy of the chosen transformation model. Ideally, a mechanical model which accounts for the patient specific tissue properties should be used. This requires, however, an identification of the tissue spatial distribution as well as the knowledge of the patient specific tissue properties. A more practical, but less accurate model, would be an isotropic homogeneous linear elastic model, like the elastic body spline [Davis et al., 1997].

Instead of relying on the precise localization of the fiducial markers and the accurate determination of the biomechanical properties for a given patient, we propose the following: to simulate plausible gold standard deformations using biomechanical models based on finite element methods. Here, the role of the biomechanical model is not to predict an actual deformation for a selected patient, but to simulate a deformation which could possibly have occurred for some patient. This general concept is applicable to any registration problem where misalignment was caused by patient motion as long as a plausible biomechanical model can be constructed. Realistically misaligned image pairs require simulation of the intensity changes (e.g. noise, image artifacts, acquisition protocol) or acquisition of aligned image pairs.

For DCE MR mammography, this work employed the second approach, by selecting from a large database the rare cases of pre- and post-contrast image pairs which showed no visible motion artifacts. Misaligned contrast-enhanced image pairs with known transformations

\(^{12}\text{Pairs of points with specially selected features which are visible in both images.}\)
can be created and hence target registration errors can be derived. The aim of this novel validation method is to provide a gold standard of sufficient realism and accuracy to allow verification of registration accuracy. Chapters 5 and 6 discuss this novel method in depth.

2.8 Summary

A large number of registration algorithms for DCE MR mammography have been reported since 1995, see Table 2.4 and Table 2.5. The general trend over those years was a shift from rigid to non-rigid transformation models, from functional to statistical voxel based similarity measures and from manual to automatic methods.

An evaluation of the quantitative performance of these algorithms has, to date, been almost non-existent. The only plausible method so far has been visual assessment, but this is not possible in regions with contrast enhancement. Other performance assessments have relied on simulated deformations which are based on similar transformation models as the registration algorithm, but which are likely to bias results. Note that an analysis of the similarity measure itself does not constitute a validation.

An objective of this thesis is to improve this unsatisfactory state of validation. A novel validation method for non-rigid image registration was suggested in Section 2.7. Its development and application to DCE MR mammography is described in detail in the subsequent Chapters 5 and 6. The usefulness of the registration method to improve the classification of breast lesions is presented in Chapter 7.
Chapter 3

Review of Biomechanical Breast Models

This chapter provides background information for developing biomechanical computer models of the breast using finite element methods (FEMs). This includes overviews of continuum mechanics, material models and FEMs. The mechanical properties of breasts are discussed followed by a review of published biomechanical breast models.

3.1 Introduction

A novel validation method for intra-visit image registration was proposed in the previous chapter, see Section 2.7. This validation method is based on simulating plausible gold standard deformations using biomechanical computer models. This chapter provides background information for creating biomechanical breast models in order to validate DCE MR mammography registration. The commercial package [ANSYS Inc., 2000] was employed for the finite element modelling and hence some of its characteristics will be described in this chapter.

The process of simulating the mechanical behaviour of a deforming media begins with the creation of a mathematical model as shown in Figure 3.1. This mathematical model is unavoidably an idealization of the true physical problem, since it is impossible to reproduce the full complexity which is present in reality. Various assumptions have to be made at this stage. The resulting differential equations are generally still too complex to be solved analytically. Instead, the FEM is employed to obtain a numerical approximate solution. The essence of the FEM is to partition the complicated domain into smaller regions (elements) for which the differential equations can approximately be solved.

![Figure 3.1: Outline of mechanical computer modelling process, adapted from [Bathe, 1996].](image-url)
3.1 Introduction

The chapter is structured as shown in Figure 3.2. Sections 3.2 to 3.4 describe the main components of mechanical computer models. These comprise of continuum mechanics, material models and FEMs.

- **Continuum mechanics** deals with continuous matter, including solids and fluids. The fact that matter is made of atoms and that it commonly has a heterogeneous microstructure is ignored. The behaviour of a very large number of molecules is instead approximated by a few quantities, such as energy and density. Continuum mechanics provides the fundamental concepts and definitions of this macroscopic mechanical behaviour. These cover the general physical laws which characterize kinematics, stresses and balance principles. These laws hold for all continuum bodies.

- **Material models** define the material response functions of individual materials. These equations, termed constitutive relations, approximate the observed macroscopic behaviour of real material under the specific conditions of interest. For solids these are given by stress-strain relationships.

- **FEMs** are a numerical technique for solving the differential equations provided by a mathematical model. In the case of mechanical models, the mathematical model is based on the fundamental equations from continuum mechanics and the constitutive relations of the material at hand.

Published work on the mechanical properties of breast tissue and on biomechanical breast models are reviewed in Section 3.5 and Section 3.6 respectively. The chapter concludes with a summary in Section 3.7.

---

1Kinematics is the study of motion and deformation without regard for the cause.
3.2 Continuum Mechanics

The aim of this section is to provide a brief description of the relevant continuum mechanics theory for modeling breast deformations. The stated definitions and derivations are from [Holzapfel, 2001]. A more detailed account can be found in [Holzapfel, 2001] and [Marsden and Hughes, 1994]. Although there was no need for the author to implement any of these textbook equations due to the use of [ANSYS Inc., 2000], the section was included for completeness and reference purposes. The reader familiar with continuum mechanics may proceed with Section 3.3.

Instead of predicting the motion of a body from the interactions of its individual molecules, continuum mechanics approximates the behaviour of a very large number of molecules by a few quantities, such as density and velocity. This macroscopic system will not be exact, but it has the advantage of being computationally tractable while being accurate enough for many modelling problems [Holzapfel, 2001].

In this section, the main concepts of continuum mechanics will be described, including kinematics, stress, equilibrium equations and the principle of virtual displacements.

3.2.1 Kinematics

A fundamental assumption of continuum mechanics is that a body can be viewed as having a (piecewise) continuous distribution of matter in space and time. Such a continuum body is then imagined as being composed of a set of (continuous) particles, where each particle consists of a large number of molecules while still being small.

Configurations: Let \( B \) be a continuum body occupying the three-dimensional region \( \Omega \) at time \( t \) in a stationary Cartesian coordinate system as shown in Figure 3.3. The region occupied by \( B \) at the initial time (\( \Omega_0 \)) is called the reference configuration while \( \Omega \) is referred to as the current configuration. To discriminate between these configurations, uppercase letters will be used for entities and indices of the reference configuration (e.g. \( X_A \)) while lowercase letters denote entities and indices of the current configuration (e.g. \( x_a \)).

Motion: Let \( \mathbf{X} = (X_1, X_2, X_3) \in \Omega_0 \) be the position of particle \( P \) in the body \( B \) at time 0. The position of this particle at time \( t \) is then described by

\[
\mathbf{x} = \Phi(\mathbf{X}, t) \quad \text{or} \quad x_a = \Phi_a(X_1, X_2, X_3, t) \quad \text{for} \quad a = 1, 2, 3, \tag{3.1}
\]

where the vector field \( \Phi \) is called the motion of body \( B \). For a continuous and one-to-one motion \( \Phi \), the inverse motion \( \Phi^{-1} \) exists with \( \mathbf{X} = \Phi^{-1}(\mathbf{x}, t) \). Motion \( \Phi \) can change the shape, the position and the orientation of a body. A body which is able to change its shape is said to be deformable.
Displacement: The positions $X$ and $x$ are related by

$$ U(X,t) = x(X,t) - X \quad \text{or} \quad U_A = x_a - X_A \quad \text{for } A = a \in \{1, 2, 3\}, \quad (3.2) $$

where $U(X,t)$ denotes the displacement field with respect to the reference configuration $\Omega_0$. By definition, displacements are independent of the geometrical configuration, i.e. $u(x,t) = x - X(x,t)$, and hence $U(X,t) = u(x,t)$.

Deformation Gradient Tensor: A deformation is quantified by its deformation gradient tensor $F$, which carries the line segment $dX$ into $dx$ by

$$ dx = F(X,t)dX \quad \text{or} \quad dx_a = \sum_{B=1}^{3} F_{aB}dX_B \quad \text{for } a = 1, 2, 3. \quad (3.3) $$

The deformation gradient tensor $F$ is defined as

$$ F(X,t) = \frac{\partial \Phi}{\partial X} \quad \text{or} \quad F_{aB} = \frac{\partial \Phi_a}{\partial X_B} = \frac{\partial x_a}{\partial X_B} = \delta_{aB} + \frac{\partial u_a}{\partial X_B} \quad \text{for } a, B \in \{1, 2, 3\}, \quad (3.4) $$

where $\delta_{aB}$ denotes the Kronecker delta. Figure 3.3 illustrates the change of line segment $dX$ due to motion $\Phi$ into line segment $dx$. If the inverse motion $\Phi^{-1}$ exists, then the inverse of $F$, carrying $dx$ into $dX$, can be expressed as

$$ F^{-1}(x,t) = \frac{\partial \Phi^{-1}}{\partial x} \quad \text{or} \quad F^{-1}_{ab} = \frac{\partial \Phi_{ab}^{-1}}{\partial x} = \frac{\partial X_a}{\partial x_b} = \delta_{Ab} - \frac{\partial u_A}{\partial x_b} \quad \text{for } b, A \in \{1, 2, 3\}. \quad (3.5) $$
Rotation and Stretch Tensor: The deformation gradient tensor \( F \) can be uniquely decomposed into a pure rotation and a pure stretch via
\[
F = RU = vR \quad \text{or} \quad F_{AB} = \sum_{K=1}^{3} R_{aK} U_{KB} = \sum_{k=1}^{3} v_{ak} R_{kB} \quad \text{for} \; a, B \in \{1, 2, 3\},
\]
where \( R \) is an orthogonal tensor with \( \det(R) = 1 \), called the rotation tensor, and \( U \) (resp. \( v \)) is a positive definite symmetric tensor termed the right (resp. left) stretch tensor.

Deformation Tensor: In order to ignore rotation \( R \), the right Cauchy-Green deformation tensor \( C = F^T F = U^T R^T R U = U^T U = U^2 \) and the left Cauchy-Green deformation tensor \( d = FF^T = vRR^T v^T = vv^T = v^2 \) are introduced\(^2\). \( C \) and \( d \) are symmetric and positive definite by definition, and their components are
\[
C_{AB} = \sum_{k=1}^{3} F_{kA} F_{kB} = \sum_{k=1}^{3} \frac{\partial x_k}{\partial X_A} \frac{\partial x_k}{\partial X_B} \quad \text{for} \; A, B \in \{1, 2, 3\},
\]
\[
d_{ab} = \sum_{k=1}^{3} F_{aK} F_{bK} = \sum_{K=1}^{3} \frac{\partial x_a}{\partial X_K} \frac{\partial x_b}{\partial X_K} \quad \text{for} \; a, b \in \{1, 2, 3\}.
\]

Principle Stretches and Strain Tensor: The eigenvalues of \( U \) and \( v \) are equal, real and positive. These eigenvalues, denoted as \( \lambda_1, \lambda_2, \lambda_3 \), are called the principle stretches. The deviation of the principle stretches from unity measures the amount of strain in a deformation. Strain quantifies the change in length of a line segment caused by a deformation and hence is unitless. In the reference configuration \( \Omega_0 \), the change in squared lengths is measured by the Green-Lagrange strain tensor \( E \) which is defined by
\[
E = \frac{1}{2} (C - I) \quad \text{or} \quad E_{AB} = \frac{1}{2} \left( \sum_{k=1}^{3} F_{kA} F_{kB} - \delta_{AB} \right) \quad \text{for} \; A, B \in \{1, 2, 3\},
\]
where \( I \) denotes the identity matrix and \( 1/2 \) is a normalization factor. \( E \) can be expressed in terms of displacements and initial coordinates as
\[
E_{AB} = \frac{1}{2} \left( \frac{\partial U_A}{\partial X_B} + \frac{\partial U_B}{\partial X_A} + \sum_{K=1}^{3} \frac{\partial U_K}{\partial X_A} \frac{\partial U_K}{\partial X_B} \right) \quad \text{for} \; A, B \in \{1, 2, 3\}.
\]
The change in squared length in the current configuration \( \Omega \) is measured by the Euler-Almansi strain tensor \( e \) given by
\[
e = \frac{1}{2} (I - d^{-1}) \quad \text{or} \quad e_{ab} = \frac{1}{2} \left( \delta_{ab} - \sum_{K=1}^{3} F_{K a}^{-1} F_{K b}^{-1} \right) \quad \text{for} \; a, b \in \{1, 2, 3\},
\]
with components
\[
e_{ab} = \frac{1}{2} \left( \frac{\partial u_a}{\partial x_b} + \frac{\partial u_b}{\partial x_a} - \sum_{k=1}^{3} \frac{\partial u_k}{\partial x_a} \frac{\partial u_k}{\partial x_b} \right) \quad \text{for} \; a, b \in \{1, 2, 3\}.
\]
\(^2\)This holds since \( R^T R = I \) for an orthogonal tensor \( R \) and \( U^T = U \) for a symmetric tensor \( U \).
For small displacements, the last term in equation (3.10) and (3.12) is much smaller than the other two. \( E \) and \( e \) can then be approximated with the infinitesimal strain tensor

\[
\varepsilon_{ab} = \frac{1}{2} \left( \frac{\partial u_a}{\partial x_b} + \frac{\partial u_b}{\partial x_a} \right) \quad \text{for} \ a, b \in \{1, 2, 3\}. \tag{3.13}
\]

All essential definitions for describing the motion of a body and its deformation are now introduced. What may have caused the motion is discussed next.

### 3.2.2 Stress

Two kinds of external forces may act on a body, namely surface force \( f \) (per unit area) or body force \( b \) (per unit mass) as shown in Figure 3.4a. Surface forces are distributed over the surface of a body, as for example pressure of one body on another. Body forces are distributed over the volume of a body, as for example gravitational forces or magnetic forces. Recall that force is measured in Newtons (\( N = \text{kg m/s}^2 \)).

Stress is defined as the force per unit area acting on an infinitesimally small plane surface. Stress is measured in Pascal (\( Pa = N/m^2 \)). The force and area quantities can be defined in the reference or the current configuration. For large deformations, it is therefore important to discriminate between three stress formulations.

The Cauchy stress tensor \( \sigma \) represents the force measured per unit area of the current configuration acting on a surface element of the current configuration. The traction vector (surface forces per unit area) is given by

\[
t(x, t, n) = \sigma(x, t)n(x) \quad \text{or} \quad t_a = \sum_{b=1}^{3} \sigma_{ab} n_b \quad \text{for} \ a \in \{1, 2, 3\}, \tag{3.14}
\]

where \( n \) is the unit normal vector of the surface element of the current configuration. \( \sigma_{ab} \) can be interpreted as the \( a \)th component of the traction vector acting on the surface \( b \) as shown in Figure 3.4b. \( \sigma_{aa} \) denotes the normal stress and \( \sigma_{ab} \), with \( a \neq b \), the shear stress. \( \sigma \) is symmetric for non-polar materials\(^3\), like biological tissue. The area in the current configuration is generally unknown for large deformations.

The first Piola-Kirchhoff stress tensor \( P \) represents the force measured per unit area in the reference configuration acting on the current configuration. The associated traction vector is defined as

\[
T(x, t, X, N) = P(x, t, X)N(X) \quad \text{or} \quad T_a = \sum_{B=1}^{3} P_{aB} N_B \quad \text{for} \ a \in \{1, 2, 3\}, \tag{3.15}
\]

\(^3\)Polar materials have permanent dipoles and hence introduce a local source of angular momentum.
Figure 3.4: a) Illustration of external forces acting on the current configuration with \( df = td_s \). b) Cauchy stress components of traction vector \( t \) acting on a cube with sides parallel to coordinate axes. The figures were adapted from [Holzapfel, 2001].

where \( \mathbf{N} \) represents the unit normal vector of the surface element of the reference configuration. \( \mathbf{P} \) is also called engineering or nominal stress and it is the stress often measured in experiments.

The second Piola-Kirchhoff stress tensor \( \mathbf{S} \) represents the force measured per unit area in the reference configuration acting on the reference configuration. \( \mathbf{S} \) cannot be measured since force only acts on the current configuration. It is, however, a useful theoretic measure for expressing stress in the reference configuration.

The three stress tensors can be related to each other using the relationship of traction in the reference and current configuration, \( df = td_s = TdS \) and \( ds = JF^{-T}dS \), with volume ratio \( J = \det(F) = \lambda_1\lambda_2\lambda_3 \). For \( a, b, A, B \in \{1, 2, 3\} \), the relationships are

\[
\begin{align*}
\mathbf{P} &= JF^{-1}\mathbf{\sigma} \\
P_{aB} &= J \sum_{b=1}^{3} \frac{\partial X_B}{\partial x_b} \sigma_{ab} \\
S_{AB} &= \sum_{a=1}^{3} P_{aB} \frac{\partial X_A}{\partial x_a} \\
S &= \mathbf{FP}^{-T} \\
S &= JF^{-1} \mathbf{F} \mathbf{S} \mathbf{F}^{-T}
\end{align*}
\]

(3.16)

\[
\begin{align*}
\mathbf{P} &= \mathbf{SF}^T \\
P_{aB} &= \sum_{A=1}^{3} S_{AB} \frac{\partial x_a}{\partial X_A} \\
\sigma_{ab} &= \frac{1}{J} \sum_{B=1}^{3} \frac{\partial x_b}{\partial X_B} P_{aB} \\
\mathbf{\sigma} &= \frac{1}{J} \mathbf{FP} \\
\mathbf{\sigma} &= \frac{1}{J} \mathbf{FSF}^T
\end{align*}
\]

(3.17)

The second Piola-Kirchhoff stress tensor \( \mathbf{S} \) is symmetric when the Cauchy stress tensor \( \mathbf{\sigma} \) is symmetric (from equation (3.16)).

This section introduced the notion of stress, which is responsible for the deformation of
3.2 Continuum Mechanics 97

materials. The necessary and sufficient conditions for a body to be in balance are described next.

3.2.3 Equilibrium Equations

A body is in equilibrium when the resultant force acting on it is zero. For static problems, where the load is applied very slowly such that the data is independent of time, the equilibrium equation for the current configuration is

\[
\int_{\partial \Omega} \mathbf{t} \, ds + \int_{\Omega} \rho \mathbf{b} \, dv = 0, \tag{3.18}
\]

where \( \mathbf{t} \) is the traction vector, \( \mathbf{b} \) is the body force and \( \rho \) is the density of \( \Omega \). Equation (3.18) can be expressed with volume integrals by applying the divergence theorem \((\int_{\partial \Omega} \mathbf{t} \, ds = \int_{\Omega} \text{div} \mathbf{\sigma} \, dv)\). Since (3.18) must hold for any arbitrary volume \( v \), it can be shown that this implies the so-called Cauchy’s equation of equilibrium:

\[
\text{div}\mathbf{\sigma} + \rho \mathbf{b} = 0 \quad \text{or} \quad \sum_{b=1}^{3} \frac{\partial \sigma_{ab}}{\partial x_{b}} + \rho b_{a} = 0 \quad \text{for} \ a \in \{1,2,3\}. \tag{3.19}
\]

Similarly, for the reference configuration this results in

\[
\text{Div}\mathbf{\Pi} + \rho_{0} \mathbf{B} = 0 \quad \text{or} \quad \sum_{B=1}^{3} \frac{\partial \Pi_{aB}}{\partial X_{B}} + \rho_{0} B_{a} = 0 \quad \text{for} \ a \in \{1,2,3\}, \tag{3.20}
\]

where \( \rho_{0} \) and \( \mathbf{B} \) are the density and the body force of \( \Omega_{0} \), respectively.

To solve the static problem, the equilibrium equations (3.19) and (3.20) need to be satisfied for every elementary volume, subject to boundary conditions, which are described next.

3.2.4 Boundary Conditions

Two types of boundary conditions can be distinguished, namely displacements \( \mathbf{u} \) and surface tractions \( \mathbf{t} \). The overbars indicate that these are prescribed values. These boundary conditions are mutually exclusive and hence partition the surface into regions where only one type of boundary condition is applied to. For the current configuration with boundary \( \partial \Omega \), this can be summarized as

\[
\mathbf{u}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}) \quad \text{on} \quad \partial \Omega^{u}, \quad \mathbf{t}(\mathbf{x}, t, \mathbf{n}) = \mathbf{\sigma}(\mathbf{x}, t) \mathbf{n} = \mathbf{t}(\mathbf{x}) \quad \text{on} \quad \partial \Omega^{t} \quad \text{with} \quad \partial \Omega^{u} \subset \partial \Omega, \quad \partial \Omega^{t} \subset \partial \Omega, \quad \partial \Omega^{u} \cap \partial \Omega^{t} = \{\}, \quad \partial \Omega^{u} \cup \partial \Omega^{t} = \partial \Omega. \tag{3.21}
\]

The boundary conditions for the reference configuration are given by

\[
\mathbf{U}(\mathbf{X}, t) = \mathbf{U}(\mathbf{X}) \quad \text{on} \quad \partial \Omega_{0}^{U}, \quad \mathbf{T}(\mathbf{X}, t, \mathbf{N}) = \mathbf{P}(\mathbf{X}, t) \mathbf{N} = \mathbf{T}(\mathbf{X}) \quad \text{on} \quad \partial \Omega_{0}^{T} \quad \text{with} \quad \partial \Omega_{0}^{U} \subset \partial \Omega_{0}, \quad \partial \Omega_{0}^{T} \subset \partial \Omega_{0}, \quad \partial \Omega_{0}^{U} \cap \partial \Omega_{0}^{T} = \{\}, \quad \partial \Omega_{0}^{U} \cup \partial \Omega_{0}^{T} = \partial \Omega_{0}. \tag{3.22}
\]
3.2 Continuum Mechanics

The boundary-value problem is then stated by (3.19) subject to (3.21), or by (3.20) subject to (3.22). Generally, it is difficult to solve the boundary-value problem in this differential formulation. Instead, a variational approach is employed, which is discussed below.

3.2.5 Principle of Virtual Displacements

The boundary-value problem is difficult to solve since the equilibrium equations (3.19) and (3.20) apply to every elementary volume. An alternative solution approach is the variational formulation, where the problem is expressed through minimizing a scalar cost function, as previously discussed in Section 2.2.3.1. The potential energy, which is given by the difference between strain energy of the system and the potential energy of the loads, is an example of such a scalar cost function.

The principle of virtual displacements, or principle of virtual work, is employed to derive the variational formulation for the analysis of solids and structure. The general approach consists of: weighting the differential equations by suitable test functions; integrating the resulting equations; and substituting the boundary conditions. This general approach can be applied to any analysis problem where the governing differential equations are known. Its application to the given boundary-value problem is described with the following steps [Holzapfel, 2001; Bathe, 1996]:

Step 1 In the case of virtual displacements, the test functions are small displacements. The qualifier "virtual" is used to emphasize that the physical quantities are hypothetical. Suppose a very small, arbitrary displacement \( \delta u \), which is compatible with the displacement boundary conditions (i.e. \( \delta u(x) = 0 \) if \( x \in \partial \Omega^u \)), is applied to the current configuration in equilibrium. The variational form of (3.19) is then given by

\[
\int_{\Omega} (\text{div} \sigma + \rho \text{b}) \cdot \delta u \ dv = 0. \tag{3.23}
\]

Step 2 Equation (3.23) is integrated using \( \text{div} \sigma \cdot \delta u = \text{div}(\sigma \delta u) - \sigma : \text{grad} \delta u \) (product rule), \( \delta e = \text{sym} \text{grad} \delta u \) and \( \sigma : \text{grad} \delta u = \sigma : \text{sym} \text{grad} \delta u \) (since \( \sigma \) is symmetric) to yield

\[
\int_{\Omega} \left( \text{div}(\sigma \delta u) - \sigma : \delta e + \rho \text{b} \cdot \delta u \right) \ dv = 0. \tag{3.24}
\]

Applying the divergence theorem \( \int_{\Omega} \text{div}(\sigma \delta u) \ dv = \int_{\partial \Omega} \sigma \delta u \cdot n \ ds \) results in

\[
\int_{\Omega} (\sigma : \delta e - \rho \text{b} \cdot \delta u) \ dv = \int_{\partial \Omega} \sigma \delta u \cdot n \ ds. \tag{3.25}
\]

Step 3 Taking account of the traction boundary conditions from (3.21), and of the fact that \( \delta u \) is zero on \( \partial \Omega^u \), this leads to

\[
\int_{\Omega} \sigma : \delta e \ dv = \int_{\Omega} \rho \text{b} \cdot \delta u \ dv + \int_{\partial \Omega^t} \tilde{t} \cdot \delta u \ ds \quad \text{or}
\]

\[
\int_{\Omega} \sigma : \delta e \ dv = \int_{\Omega} \rho \text{b} \cdot \delta u \ dv + \int_{\partial \Omega^t} \tilde{t} \cdot \delta u \ ds.
\]
\[ \int_{\Omega} \sum_{a,b=1}^{3} \sigma_{ab} \delta e_{ab} dV = \int_{\Omega} \rho \sum_{a=1}^{3} b_a \delta u_a dV + \int_{\partial \Omega} \sum_{a=1}^{3} \overline{t}_a \delta u_a ds. \] (3.26)

The stress \( \sigma \) working along the virtual strain \( \delta e \) creates the internal virtual work \( \delta W_{int} \). The work done by the loads (body force \( b \) and surface traction \( \overline{t} \)) along the virtual displacement \( \delta u \) defines the external virtual work \( \delta W_{ext} \). This variational formulation states that the equilibrium of a body is reached when internal and external work (instead of force) balance.

For the reference configuration, the virtual work equation can be expressed with respect to the first Piola-Kirchhoff stress tensor \( P \) via

\[ \int_{\Omega_0} P : \delta F dV = \int_{\Omega_0} \rho_0 B : \delta U dV + \int_{\partial \Omega_0^T} \overline{T} : \delta U dS \quad \text{or} \]
\[ \int_{\Omega_0} \sum_{a,B=1}^{3} P_{aB} \delta F_{aB} dV = \int_{\Omega_0} \rho_0 \sum_{a=1}^{3} B_a \delta U_a dV + \int_{\partial \Omega_0^T} \sum_{a=1}^{3} \overline{T}_a \delta U_a dS, \] (3.27)

or with respect to the second Piola-Kirchhoff stress tensor \( S \)

\[ \int_{\Omega_0} S : \delta E dV = \int_{\Omega_0} \rho_0 B : \delta U dV + \int_{\partial \Omega_0^T} \overline{T} : \delta U dS \quad \text{or} \]
\[ \int_{\Omega_0} \sum_{A,B=1}^{3} S_{AB} \delta E_{AB} dV = \int_{\Omega_0} \rho_0 \sum_{a=1}^{3} B_a \delta U_a dV + \int_{\partial \Omega_0^T} \sum_{a=1}^{3} \overline{T}_a \delta U_a dS. \] (3.28)

### 3.2.6 Summary

This section described the basic principles of continuum mechanics. Determining the motion for static scenarios was formulated as a boundary-value problem, see Section 3.2.3 and 3.2.4. This differential formulation is generally hard to solve and hence was transformed into a variational approach by means of the principle of virtual displacements.

The boundary-value problem and its variational formulation are general in the sense that they apply to any material. To solve the virtual work equations (3.26)-(3.28), the relationship between stress and strain needs to be known. This constitutive relationship, termed material model, determines the specific behaviour of the material under consideration. Breast specific material models are described in detail in Section 3.5 while material models are discussed in general in the next section.

In principle, the virtual work equations may be solved for the unknown displacement field \( u \), subject to the boundary constraints. In practise, however, a general solution is rarely possible. Instead, the solution is approximated by geometrical discretization of the continuous medium via the FEM as described in Section 3.4.
<table>
<thead>
<tr>
<th>Material model</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elastic</td>
<td>Stress is a function of strain only: ( \sigma_{ab} = \sum_{a,b} z_{abcd} \epsilon_{cd} ), with elasticity tensor ( z_{abcd} ), see (3.30). Same loading and unloading behaviour. Linear elastic behaviour if ( z_{abcd} ) is constant. Non-linear elastic behaviour if ( z_{abcd} ) varies as a function of strain.</td>
</tr>
<tr>
<td>Hyperelastic</td>
<td>Stress is calculated from a strain energy function ( \Psi ), see (3.32).</td>
</tr>
<tr>
<td>Elastoplastic</td>
<td>Elastic behaviour until yield, then irreversible deformations.</td>
</tr>
<tr>
<td>Viscoelastic</td>
<td>Time effect of increasing strain under constant load (creep), decreasing stress under constant deformations (relaxation), or different loading and unloading behaviour (hysteresis).</td>
</tr>
<tr>
<td>Viscoplasticity</td>
<td>Time-dependent inelastic strains; rate effects are included.</td>
</tr>
</tbody>
</table>

Table 3.1: Basic properties of common material model.

### 3.3 Material Models

Material models describe the relationship between stress and strain. Table 3.1 lists the main properties of common materials types. Breast deformations are generally simulated using elastic or hyperelastic models, which are addressed next. The definitions and derivations in this section are from [Holzapfel, 2001; Mase, 1970; eFunda, 2002; Bathe, 1996; ANSYS Inc., 2003].

#### 3.3.1 Elastic Materials

A material is called elastic if the relationship between stress and strain is independent of any history. Elastic materials recover completely to the undeformed state after removal of the applied loads. Provided the stresses are small enough, most materials are elastic. In the reference configuration, the constitutive law of elastic models can be described by

\[
S = Z : E
\]

where \( Z \) denotes the elasticity tensor defined as

\[
Z = \frac{\partial S(E)}{\partial E} \quad \text{or} \quad Z_{ABCD} = \frac{\partial S_{AB}}{\partial E_{CD}} \quad \text{for} \ A, B, C, D \in \{1, 2, 3\}.
\]

In the current configuration, the stress-strain relationship is given by

\[
\sigma = z : \epsilon,
\]

where the elasticity tensor \( z \) is defined by

\[
z_{abcd} = J^{-1} F_{aA} F_{bB} F_{cC} F_{dD} Z_{ABCD}.
\]

Note that \( Z \) and \( z \) are 4th order tensors and hence have \( 3^4 = 81 \) components. The symmetry of both, the stress tensor \( (\epsilon_{cd} = \epsilon_{dc}) \) and the strain tensor \( (\sigma_{ab} = \sigma_{ba}) \), implies that \( z_{abcd} = z_{bacd} = z_{abdc} \) and hence only 36 components of \( z \) are independent [Mase, 1970]. For
hyperelastic materials, this reduces to 21 independent components since $z = z^T$, see (3.33). Orthotropic materials possess three mutually perpendicular planes of elastic symmetry and have 12 independent components, or 9 if $z = z^T$.

Transverse isotropic materials have the same elastic properties for two of three mutually perpendicular planes, and can be described by 5 independent elastic constants. Assuming the same properties in the $x_1$-$x_2$ plane, the strain and stress relationship can be described by [eFunda, 2002]

$$
e_{11} = \frac{1}{E_p} \sigma_{11} - \frac{v_p}{E_p} \sigma_{22} - \frac{v_{3p}}{E_3} \sigma_{33}, \quad e_{22} = -\frac{v_p}{E_p} \sigma_{11} + \frac{1}{E_p} \sigma_{22} - \frac{v_{3p}}{E_3} \sigma_{33}, \quad e_{33} = -\frac{v_{3p}}{E_p} \sigma_{11} - \frac{v_{3p}}{E_p} \sigma_{22} + \frac{1}{E_3} \sigma_{33},
$$

$$
e_{23} = \frac{1}{2\mu_{3p}} \sigma_{23}, \quad e_{31} = \frac{1}{2\mu_{3p}} \sigma_{31},
$$

(3.31)

where $E_p$ and $v_p$ represent the Young's modulus and the Poisson's ratio in the $x_1$-$x_2$ plane, $E_3$ and $v_{3p}$ denote the Young's modulus and the Poisson's ratio in the $x_3$ direction, $\mu_{3p}$ stands for the shear modulus in the $x_3$ direction. Symmetry implies that $v_{3p}/E_p = v_{3p}/E_3$. Transverse isotropic materials are completely defined by $v$, $E_3$, $E_p$ and $\mu_{3p}$, if the induced volume change is independent of the applied stress direction. $v$ represents in this case the Poisson's ratio of isotropic materials. The Poisson's ratios of the transverse isotropic materials were then derived as follows by the author. For small strains, the volume change due to an uniaxial stress applications of $\sigma_{ii}$ can be approximated by $e_{ii} + e_{jj} + e_{kk} = e_{ii}(1 - 2v)$ for $i, j, k \in \{1, 2, 3\}$ and $i \neq j \neq k$. Uniaxial stress applications of $\sigma_{11}$ and the volume change approximation lead to $(1/E_p - v_p/E_p - v_{3p}/E_p)\sigma_{11} = e_{11}(1 - 2v)$, and hence $v_p + v_{3p} = 2v$.

Similarly, uniaxial application of $\sigma_{33}$ and the volume change approximation result in $v_{3p} = v$.

Finally, symmetry implies that $v_{3p} = v/r_{3p}$ and $v_p = 2v(1 - 0.5/r_{3p})$ where $r_{3p} = E_3/E_p$.

Isotropic materials are elastically equivalent in all directions and have just two independent components. These are called the Lamé constants $\lambda$ and $\mu$. The elasticity tensor is then given by $z_{abcd} = \lambda \delta_{a\delta} \delta_{cd} + \mu (\delta_{ac}\delta_{bd} + \delta_{ad}\delta_{bc})$ [Bathe, 1996]. The Lamé constants are related to Young’s modulus $E$ and Poisson’s ratio $\nu$ by $\lambda = E\nu/[(1 + \nu)(1 - 2\nu)]$ and $\mu = E/[2(1 + \nu)]$. For an uniaxial stress scenario in direction $x_1$, $E$ and $v$ are given by the relationship $\sigma_{11} = E e_{11}$ and $e_{33} = e_{22} = -v e_{11}$.

### 3.3.2 Hyperelastic Materials

An elastic material is said to be hyperelastic if there exists a strain energy function $\Psi$ which completely determines the stress to strain relation by [Holzapfel, 2001]

$$
S = 2\frac{\partial \Psi(C)}{\partial C}, \quad P = 2F \frac{\partial \Psi(C)}{\partial C}, \quad \sigma = 2\frac{1}{J} F \frac{\partial \Psi(C)}{\partial C} F^T.
$$

(3.32)
Hyperelastic materials have a symmetrical elasticity tensor, i.e. $Z = Z^T$, since

$$Z = 4 \frac{\partial^2 \Psi(C)}{\partial C \partial C} \quad \text{or} \quad Z_{ABCD} = 4 \frac{\partial \Psi}{\partial C_{AB} \partial C_{CD}} = 4 \frac{\partial \Psi}{\partial C_{CD} \partial C_{AB}} = Z_{CDAB}. \quad (3.33)$$

Symmetry of $Z$ implies that a strain energy function $\Psi$ exists and hence is also used as a definition for hyperelastic materials. Hyperelastic models are employed to describe materials which can experience large elastic strains that are recoverable, like rubber. An important subclass are isotropic hyperelastic materials, which are described next.

**Isotropic Hyperelastic Materials:** The strain energy function $\Psi$ of isotropic materials can be expressed in terms of the three invariants $I_1, I_2$ and $I_3$. These invariants are the coefficient of the characteristic polynomial for $C$, i.e. $-\lambda^3 + I_1 \lambda^2 - I_2 \lambda + I_3 = 0$, and are given by

$$I_1(C) = \text{tr}(C) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2,$$
$$I_2(C) = \frac{1}{2}[(\text{tr}(C))^2 - \text{tr}(C^2)] = \lambda_1^2 \lambda_2^2 + \lambda_2^2 \lambda_3^2 + \lambda_3^2 \lambda_1^2,$$
$$I_3(C) = \text{det}(C) = \lambda_1 \lambda_2 \lambda_3,$$

where $\text{tr}(C)$ and $\text{det}(C)$ denote the trace and the determinant of $C$ respectively, and $\lambda_1, \lambda_2$ and $\lambda_3$ are the principle stretches as defined on page 94. Note that the three invariants are not affected by coordinate rotations.

For isotropic material, the strain energy function $\Psi(C)$ can be regarded as a function of the three principle stretches, i.e. $\Psi(\lambda_1, \lambda_2, \lambda_3)$. Consequently, the constitutive equations (3.32) simplify to [Holzapfel, 2001]:

$$S_a = \frac{1}{\lambda_a} \frac{\partial \Psi}{\partial \lambda_a}, \quad P_a = \frac{\partial \Psi}{\partial \lambda_a}, \quad \sigma_a = \frac{1}{J} \lambda_a \frac{\partial \Psi}{\partial \lambda_a}, \quad \text{for} \ a = 1, 2, 3, \quad (3.35)$$

where $S_a, P_a$ and $\sigma_a$ denote the principle stresses.

Numerous strain-energy functions have been proposed. Two common functions are the neo-Hookean and the Mooney-Rivlin model, defined by [ANSYS Inc., 2003]

$$\text{neo-Hookean: } \Psi = \alpha_1 (I_1 - 3), \quad (3.36)$$
$$\text{2 parameter Mooney-Rivlin: } \Psi = \alpha_1 (I_1 - 3) + \alpha_2 (I_2 - 3), \quad (3.37)$$
$$\text{5 parameter Mooney-Rivlin: } \Psi = \alpha_{10} (I_1 - 3) + \alpha_{01} (I_2 - 3) + \alpha_{20} (I_1 - 3)^2 + \alpha_{11} (I_1 - 3)(I_2 - 3) + \alpha_{02} (I_2 - 3)^2. \quad (3.38)$$

**Fitting of Hyperelastic Models:** The model constants $\alpha_i$ of (3.36)-(3.38) are determined by fitting $\Psi$ to the experimental stress-strain data from mechanical deformation tests. Incompressible, isotropic hyperelastic material can be accurately characterized using the three simple independent deformation tests illustrated in Figure 3.5. To achieve accurate
3.3 Material Models

1. Uniaxial tension or equibiaxial compression,

2. Uniaxial compression or equibiaxial tension,

3. Planar tension or planar compression.

Figure 3.5: Deformation tests for incompressible, isotropic hyperelastic materials.

and stable models, test data from several modes of deformation over a wide range of strain values should be used. For breast tissue samples, only data from uniaxial compression experiments have been reported. Fitting hyperelastic models to the results of these experiments is described next.

The strain energy function $\Psi$ can be fitted to an uniaxial compression experiment in the following way [Holzapfel, 2001; ANSYS Inc., 2003]. Let $\lambda_1 = \lambda$ be the stretch in the load direction ($\sigma_1 = \sigma$) and $\lambda_2 = \lambda_3$ be the stretches in the directions that are not loaded ($\sigma_2 = \sigma_3 = 0$). Then incompressibility constraint ($\lambda_1 \lambda_2 \lambda_3 = 1$) implies that $\lambda_2 = \lambda_3 = \lambda^{-1/2}$. The invariants $I_1$ and $I_2$ from (3.34) and their derivatives are therefore

$$I_1 = \lambda^2 + 2\lambda^{-1}, \quad \frac{\partial I_1}{\partial \lambda} = 2(\lambda - \lambda^{-2}), \quad I_2 = 2\lambda + \lambda^{-2}, \quad \frac{\partial I_2}{\partial \lambda} = 2(1 - \lambda^{-3}). \quad (3.39)$$

Using (3.35) and (3.39), the constitutive equation of the neo-Hookean model (3.36) for the current configuration is given by

$$\sigma = \lambda \frac{\partial \Psi}{\partial I_1} \frac{\partial I_1}{\partial \lambda} = 2\alpha_1(\lambda^2 - \lambda^{-1}). \quad (3.40)$$

Similarly, for the 2 parameter Mooney-Rivlin model (3.37), one obtains

$$\sigma = \lambda \frac{\partial \Psi}{\partial I_1} \frac{\partial I_1}{\partial \lambda} + \lambda \frac{\partial \Psi}{\partial I_2} \frac{\partial I_2}{\partial \lambda} = 2(\lambda^2 - \lambda^{-1}) \left[ \alpha_1 + \lambda^{-1} \alpha_2 \right]. \quad (3.41)$$

While the 5 parameter Mooney-Rivlin model (3.38) has

$$\sigma = 2(\lambda^2 - \lambda^{-1}) \left[ \alpha_{10} + \alpha_{20}2(I_1 - 3) + \alpha_{11}(I_2 - 3) + \alpha_{02}2(I_2 - 3) + \alpha_{11}(I_1 - 3) \right]. \quad (3.42)$$

The constitutive equations with respect to the first and second Piola-Kirchhoff stress can be derived from (3.40)-(3.42) using $P = \lambda^{-1}\sigma$ and $S = \lambda^{-2}\sigma$, see (3.35) with $J = 1$. 
3.3 Material Models

Figure 3.6 shows the result of fitting neo-Hookean and Mooney-Rivlin models to data from a linear and an exponential stress-strain relationship using [ANSYS Inc., 2000]. The data is in all cases well matched by the 5-parameter Mooney-Rivlin model, while the neo-Hookean model provides an approximation in all cases. Fitting the 2- and 5-parameter Mooney-Rivlin model to the compression data (Figure 3.6c) leads to an unstable stress-strain relationship, i.e. $d\sigma/de \leq 0$, for strains $e \geq 0.2$ and $e \geq 0.025$, respectively. This illustrates the problem of insufficient experimental data. Biological tissues have been reported to have exponential response functions [Fung, 1993]. Whether the function shown in Figure 3.6d is correct for tensile strains remains to be proven.

![Figure 3.6](image-url)

**Figure 3.6**: Example of fitting hyperelastic material models to the stress-strain relationship of fibroglandular tissue for data (a) from a linear function [Sarvazyan et al., 1994] and (b-d) from an exponential function [Azar et al., 2002; Wellman, 1999]. The three results for the exponential function in (b-d) differ due to the assumptions which are made for the stress-strain relationship of the tensile strain. (a,b) The same behaviour for compression and tension is assumed. (c) The models are only fitted to the compression data. (d) The exponential function is assumed to extend to tensile strain.
3.3.3 Summary

This section discussed the stress-strain relationship especially for elastic and hyperelastic materials. The process of fitting hyperelastic models to the data from uniaxial compression experiments was described and shortcomings were illustrated.

For a given material model, a solution can now be sought for the variational formulation of the boundary-value problem described in the previous section. This problem can rarely be solved directly. Instead, a numerical method is employed to find an approximation.

3.4 Finite Element Method

The FEM is a numerical technique which provides an approximate solution of a boundary-value problem described by differential equations. The FEM is employed to find a displacement field \( \mathbf{u} \) which satisfies the virtual work equation for the current (3.26) or the reference (3.28) configuration. The principal idea of the FEM is to partition a body into simpler elements, for which a solution can be derived. This section is based on definitions and derivations from [Flaherty, 2005] for the geometrical discretization (Section 3.4.1) and from [Jeremić, 2004; Holzapfel, 2001; ANSYS Inc., 2003; Bathe, 1996] for the finite element formulation (Section 3.4.2-3.4.4).

3.4.1 Geometrical Discretization

A continuous domain \( \Omega \) can be geometrically discretized into a finite number of simpler smaller elements \( \Omega^{(m)} \) as illustrated in Figure 3.7a. Each element is defined by a set of points, termed nodes, and interpolation functions, called shape functions. Neighbouring elements share nodes and element boundaries as shown in 2D in Figure 3.7b.

![Figure 3.7: FEM domain discretization showing: a) partitioning of a 2D domain into triangles and b) location of nodes (●) for element \( \Omega^{(m)} \).](image)
Shape Interpolation: Functional values (displacement, strain and stress) within each element are approximated by linearly interpolating the values at the element’s nodes using the shape functions $H_i$. For example, the current displacement of point $P$ at the global position $\mathbf{x}_p = [x_{1p}, x_{2p}, x_{3p}]^T$ lying within the $k$-nodded element $m$ is then given by

$$u(\mathbf{x}_p) = \sum_{i=1}^{k} H_i^{(m)}(\mathbf{x}_p) \mathbf{u}_i^{(m)}, \quad (3.43)$$

where $\mathbf{u}_i^{(m)} = [u_{1i}^{(m)}, u_{2i}^{(m)}, u_{3i}^{(m)}]^T$ denotes the current displacement of the $i$th node of element $m$ and $H_i^{(m)}(\mathbf{x}_p)$ is the shape function, which defines the influence of node $i$ at $\mathbf{x}_p$. To simplify the solution process, the shape function $H_i$ is required to be zero at all node positions but node $i$, where it should be unity (i.e. $H_i^{(m)}(\mathbf{x}_k^{(m)}) = \delta_{ik}$ for node $k$). This ensures that a value at a node position is not influenced by the value at any other node.

Ten-noded tetrahedral elements, as depicted in Figure 3.8b, were used in this work since they were reported to be as accurate as linear hexahedral elements [Cifuentes and Kalbag, 1992; Vicentini et al., 1998] and they allow the generation of unstructured meshes directly from triangulated surfaces. Interpolation of these elements is quadratic and their shape function can be constructed by introducing natural coordinates $L_j^{(m)}(\mathbf{x}_p)$ as described subsequently.

Natural Coordinates: Let $L_j^{(m)}(\mathbf{x}_p), j = 1, 2, 3, 4$, form a redundant coordinate system on a 4-noded tetrahedron as shown in Figure 3.8a. For a point $P$, lying within tetrahedron $m$ and having global coordinates $\mathbf{x}_p = [x_{1p}, x_{2p}, x_{3p}]^T$, the natural coordinates $L_j^{(m)}(\mathbf{x}_p)$ are then given by [Flaherty, 2005]

$$L_1^{(m)}(\mathbf{x}_p) = \frac{V_{1234}^{(m)}}{V_{p234}^{(m)}}, \quad L_2^{(m)}(\mathbf{x}_p) = \frac{V_{134}^{(m)}}{V_{p134}^{(m)}}, \quad L_3^{(m)}(\mathbf{x}_p) = \frac{V_{124}^{(m)}}{V_{p124}^{(m)}}, \quad L_4^{(m)}(\mathbf{x}_p) = \frac{V_{123}^{(m)}}{V_{p123}^{(m)}}, \quad (3.44)$$

Figure 3.8: a) Node positions for a 4-noded tetrahedron and definition of natural coordinates $L_i$, redrawn from [Flaherty, 2005]. b) Node positions and node numbering for a 10-noded tetrahedron.
where $V_{ijkl}^{(m)}$ is the volume of the tetrahedron formed by point $P$ and the nodes $j, k, l$ and $V_{1234}^{(m)}$ is the volume of tetrahedron $m$ defined by

$$V_{ijkl}^{(m)} = \det \begin{bmatrix} 1 & 1 & 1 & 1 \\ x_{1p} & x_{1j}^{(m)} & x_{1k}^{(m)} & x_{1l}^{(m)} \\ x_{2p} & x_{2j}^{(m)} & x_{2k}^{(m)} & x_{2l}^{(m)} \\ x_{3p} & x_{3j}^{(m)} & x_{3k}^{(m)} & x_{3l}^{(m)} \end{bmatrix}, \quad V_{1234}^{(m)} = \det \begin{bmatrix} x_{11}^{(m)} & x_{12}^{(m)} & x_{13}^{(m)} & x_{14}^{(m)} \\ x_{21}^{(m)} & x_{22}^{(m)} & x_{23}^{(m)} & x_{24}^{(m)} \\ x_{31}^{(m)} & x_{32}^{(m)} & x_{33}^{(m)} & x_{34}^{(m)} \end{bmatrix}. \quad (3.45)$$

The natural coordinates for node $i$ are given by $L_j^{(m)}(x_i^{(m)}) = \delta_{ji}$ since $V_{ijkl}^{(m)} = 0$ if $i \in \{j, k, l\}$ and $V_{ijkl}^{(m)} = V_{1234}^{(m)}$ if $i \neq j \neq k \neq l$.

**Shape Functions:** For 4-noded tetrahedrons, the shape functions are linear and equal to the natural coordinates, i.e. $H_j^{(m)}(x_p) = L_j^{(m)}(x_p)$ and $x_{ip} = \sum_{j=1}^{4} L_j^{(m)}(x_p) x_i^{(m)}$.

Tetrahedrons with quadratic shape interpolation require an additional node on each edge of the tetrahedron, see Figure 3.8b. Shape function $H_j^{(m)}(x_p)$ can be found by employing the natural coordinates defined in (3.44), and taking care of the requirement that $H_j^{(m)}(x_k^{(m)}) = \delta_{jk}$ for node $k$. The shape function of two example nodes are deduced as follows:

- For node 1, $L_1^{(m)}(x_k^{(m)}) = 0$ for nodes $k \in \{2, 3, 4, 6, 9, 10\}$ and $L_1^{(m)}(x_k^{(m)}) = 1/2$ for nodes $k \in \{5, 7, 8\}$. Assuring that $H_1^{(m)}(x_1^{(m)}) = 1$ implies that

$$H_1^{(m)}(x) = 2L_1^{(m)}(x) \left( L_1^{(m)}(x) - 1/2 \right). \quad (3.46)$$

- For midside node 5, $L_1^{(m)}(x_k^{(m)}) = 0$ for nodes $k \in \{2, 3, 4, 6, 9, 10\}$ and $L_2^{(m)}(x_k^{(m)}) = 0$ for nodes $k \in \{1, 3, 4, 7, 8, 10\}$. $L_1^{(m)}(x_5^{(m)}) = L_2^{(m)}(x_5^{(m)}) = 1/2$ and hence

$$H_5^{(m)}(x) = 4L_1^{(m)}(x)L_2^{(m)}(x). \quad (3.47)$$

This applies similarly to the other nodes. The shape functions of the nodes of the 10-noded tetrahedral element $m$ are therefore given by

$$H_i^{(m)}(x) = \left( 2L_i^{(m)}(x) - 1 \right) L_i^{(m)}(x) \quad \text{for} \quad i \in \{1, 2, 3, 4\}$$

$$H_i^{(m)}(x) = 4L_k^{(m)}(x)L_j^{(m)}(x) \quad \text{for} \quad i \in \{5, ..., 10\}, \ i \text{ between } k \text{ and } j. \quad (3.48)$$

Having established the shape function of the employed elements, the boundary-value problem can now be expressed with respect to the nodes of the finite element model. The finite element formulation for static linear elastic cases is derived in the next section, before considering problems of static non-linear elasticity in Section 3.4.3.
3.4.2 Finite Element Formulation for Static Linear Elasticity

The mechanical problem will only be linear if its material model is linear and if the geometric configuration does not undergo large changes (i.e. small rotations, displacements and strains). For uniaxial compressions, for example, approximating the Euler-Almansi strain tensor $e_{aa}$, as defined in equation (3.12), by the infinitesimal strain tensor $e_{aa}$ (equation (3.13)) leads to an strain error $|e_{aa}/e_{aa} - 1| = 0.5e_{aa}$, since $\partial u/a/\partial x_b = 0$ if $a \neq b$, $e_{aa} = \partial u/a/\partial x_a - 0.5(\partial u/a/\partial x_a)^2$ and $e_{aa} = \partial u/a/\partial x_a$. A linear approximation of 20% for an uniaxial compression incorporates therefore a strain error of 10%, i.e. $e_{aa} = 18\%$, and must be regarded as a large deformation.

The notion that a Poisson’s ratio $v$ of 0.5 results in a volume preserving deformations is also based on the small strain assumption. Neglecting the higher order strain terms in $V(t) = V(0)(1 + e_{11})(1 - ve_{11})^2$ when calculating the change in volume, i.e. $V(t)/V(0) - 1 = (1 - 2v)e_{11}$, will lead to a 3% error for an 20% uniaxial compression of a box and $v = 0.5$. Linearization errors remain below 2% for uniaxial compressions of less than 4% strain, which seems acceptable.

For linear problems, one does not need to discriminate between reference and current configuration. All equations are therefore given in the notation of the current configuration. The shape functions (3.48) can be employed in conjunction with the virtual work equation (3.26) to create a system of linear equations with respect to the nodal displacements. Subsequent equations are formulated in terms of all nodes of the FE mesh since this leads to an effective assemblage process termed the direct stiffness method.

The virtual displacement $\delta u$ at $x$ in element $m$ is given by interpolating the virtual displacements $\delta \hat{u}$ of the FE nodes:

$$\delta u = H\delta \hat{u} \quad \text{or} \quad \delta u_a(x) = \sum_{i=1}^{N} H_i(x)\delta \hat{u}_{ai} \quad \text{for } a \in \{1, 2, 3\}, \tag{3.49}$$

where $N$ is the number of nodes in the FE mesh and $H_i = 0$ for nodes not belonging to element $m$. Similarly, the first derivative of the virtual displacements is defined as

$$\text{grad} \delta u(x) = \text{grad}H \cdot \delta \hat{u} \quad \text{or} \quad \frac{\partial \delta u_a(x)}{\partial x_b} = \sum_{i=1}^{N} \frac{\partial H_i(x)}{\partial x_b}\delta \hat{u}_{ai} \quad \text{for } a, b \in \{1, 2, 3\}. \tag{3.50}$$

Using $\delta e = \text{sym}(\text{grad} \delta u)$ and $\sigma = z : e$, the virtual work equation (3.26) can be expressed with respect to the nodal displacements of the assemblage of $M$ finite elements by [Jeremić, 2004]
\[ \sum_{m=1}^{M} \int_{\Omega^{(m)}} \text{grad} \mathbf{H} \cdot \delta \mathbf{u} : \mathbf{Z} : \text{grad} \mathbf{H} \cdot \mathbf{u} \, dv^{(m)} = \sum_{m=1}^{M} \int_{\Omega^{(m)}} \rho \mathbf{b} \cdot \mathbf{H} \cdot \delta \mathbf{u} \, dv^{(m)} + \sum_{m=1}^{M} \int_{\partial \Omega^{(m)}} \mathbf{t} \cdot \mathbf{H} \cdot \delta \mathbf{u} \, ds^{(m)} \quad \text{or} \]
\[ \sum_{m=1}^{M} \int_{\Omega^{(m)}} \sum_{a,c=1}^{3} \sum_{i=1}^{N} \frac{\partial H_i}{\partial x_b} \delta \mathbf{u}_{ai} z_{abcd} \sum_{j=1}^{N} \frac{\partial H_j}{\partial x_d} \mathbf{t}_{cj} \, dv^{(m)} = \sum_{m=1}^{M} \int_{\Omega^{(m)}} \sum_{a=1}^{3} \mathbf{t}_a H_i \, ds^{(m)} \quad \text{for } i = 1, 2, \ldots, N, \text{ and } a = 1, 2, 3. \]  

Factoring out \( \delta \mathbf{u}_{ai} \) and \( \mathbf{t}_{cj} \), and cancelling the virtual displacement \( \delta \mathbf{u}_{ai} \) on both sides, results in [Jeremić, 2004]

\[ \sum_{m=1}^{M} \left[ \int_{\Omega^{(m)}} \text{grad} \mathbf{H} : \mathbf{Z} : \text{grad} \mathbf{H} \, dv^{(m)} \right] \mathbf{u} = \sum_{m=1}^{M} \int_{\Omega^{(m)}} \rho \mathbf{b} \cdot \mathbf{H} \, dv^{(m)} + \sum_{m=1}^{M} \int_{\partial \Omega^{(m)}} \mathbf{t} \cdot \mathbf{H} \, ds^{(m)} \quad \text{or} \]
\[ \sum_{m=1}^{3} \sum_{a=1}^{M} \sum_{i=1}^{N} \left[ \int_{\Omega^{(m)}} \frac{\partial H_i}{\partial x_b} z_{abcd} \frac{\partial H_j}{\partial x_d} \, dv^{(m)} \right] \mathbf{u}_{aj} = \sum_{m=1}^{M} \int_{\partial \Omega^{(m)}} \mathbf{t}_a H_i \, ds^{(m)} \quad \text{for } i = 1, 2, \ldots, N, \text{ and } a = 1, 2, 3. \]  

Defining the element stiffness tensor \( k_{iaj}^{(m)} \), the tensor of element body force \( f_{ia}^{(m)b} \) and the tensor of element traction force \( f_{ia}^{(m)t} \) as

\[ k_{iaj}^{(m)} = \int_{\Omega^{(m)}} \frac{\partial H_i}{\partial x_b} z_{abcd} \frac{\partial H_j}{\partial x_d} \, dv^{(m)}, \]
\[ f_{ia}^{(m)b} = \int_{\Omega^{(m)}} \rho \mathbf{b}_a H_i \, dv^{(m)}, \]
\[ f_{ia}^{(m)t} = \int_{\partial \Omega^{(m)}} \mathbf{t}_a H_i \, ds^{(m)}, \]  

equation (3.52) can be expressed as

\[ \sum_{c=1}^{3} \sum_{j=1}^{N} \sum_{m=1}^{M} k_{iaj}^{(m)} \mathbf{u}_{cj} = \sum_{m=1}^{M} f_{ia}^{(m)b} + \sum_{m=1}^{M} f_{ia}^{(m)t} \quad \text{for } i = 1, 2, \ldots, N \text{ and } a = 1, 2, 3. \]  

Summing over all element tensors results in

\[ \sum_{c=1}^{3} \sum_{j=1}^{N} K_{iaj} \mathbf{u}_{cj} = f_{ia} \quad \text{for } i = 1, 2, \ldots, N \text{ and } a = 1, 2, 3, \]  

where \( K_{iaj} = \sum_{m=1}^{M} k_{iaj}^{(m)} \) is the system stiffness tensor, \( \mathbf{u}_{cj} \) is the tensor of unknown generalized nodal displacements and \( f_{ia} = \sum_{m=1}^{M} f_{ia}^{(m)b} + \sum_{m=1}^{M} f_{ia}^{(m)t} \) is the system load tensor.

After rearranging the tensors from equation (3.55) consistently for the 3 degrees of freedom per node, and taking account of displacement boundary conditions on \( \partial \Omega^u \), this can be
written in matrix and vector notation as

$$K\hat{u} = f \text{ or } \sum_{p=1}^{3N} K_{pq} \hat{u}_p = f_q \text{ for } x_p \text{ not on } \partial \Omega^u, \ q = 1, 2, \ldots, 3N,$$

$$\hat{u}_p = \bar{u}(x_p) \text{ for } x_p \text{ on } \partial \Omega^u,$$

(3.56)

$K$ and $f$ are called the stiffness matrix and the force vector, respectively. $K$ is a known matrix, since it is a function of the material properties, the solid’s geometry, the interpolation functions and the nodal positions. Similarly, $f$ is a function of the known traction boundary conditions, body forces, interpolation functions and nodal positions. Equation (3.56) is a system of $3N$ linear equations for $3N$ nodal displacements.

[ANSYS Inc., 2000], the commercial FEM package used in this work, provides several methods for solving equation (3.56) using either a direct elimination process or an iterative approach. Two direct solvers are available in ANSYS, namely the sparse direct solver and the frontal solver. The sparse direct solver employs LU or Cholesky decomposition [Press et al., 1992] to solve (3.56). It takes advantage of the many zero entries in $K$ by storing only nonzero entries, and by reordering the equations such that the number of zero entries which must become nonzero during the decomposition is kept to a minimum. The frontal solver employs Gauss elimination [Press et al., 1992], with the difference that only equations which are actually required for the elimination of a specific degree of freedom are assembled. Iterative solvers typically involve an initial guess $\hat{u}_0$ for $\hat{u}$. In further iterations, $\hat{u}_{i+1}$ is calculated from $K, f - K\hat{u}$ and $\hat{u}$ from one or two previous iterations. The solution generally converges to within a specified tolerance after a finite number of iterations. [ANSYS Inc., 2000] provides several iterative solvers including the Jacobi conjugate gradient solver, the preconditioned conjugate gradient solver, the incomplete Cholesky conjugate gradient solver. Generally, iterative solvers are more efficient but less robust than direct solvers.

This section described the finite element formulation for static linear elasticity problems, while the next section addresses its non-linear counterpart.
3.4 Finite Element Method

3.4.3 Finite Element Formulation for Static Non-Linear Elasticity

The virtual work equations (3.26) to (3.28) are typically non-linear with respect to the unknown displacement field $u$ due to kinematic non-linear effects (large displacements, large rotations and large strains) and non-linear material properties, and hence cannot be solved directly. Instead, an approximate solution can be obtained by referring all variables to the previously calculated equilibrium configuration, and linearizing the virtual work equation.

Linearization of the virtual work equation $\delta W(u, \delta u) = \delta W_{\text{ext}} - \delta W_{\text{int}} = 0$ is based on a first order Taylor series expansion [Holzapfel, 2001]

$$\delta W(u, \delta u, \Delta u) = \delta W(u, \delta u) + D\Delta u \delta W(u, \delta u) + o(\Delta u, \delta u),$$  \hspace{1cm} (3.57)

where $D\Delta u \delta W(u, \delta u)$ denotes the linear change of function $\delta W(u, \delta u)$ due to displacement field increment $\Delta u$ at $u$, and $o(\Delta u, \delta u)$ represents a small error, which decreases to zero faster than $\Delta u$ does, which henceforth is neglected.

Assuming that the external loads ($b$ and $t$) of the virtual work equation (3.26) are independent of the deformation of the body (i.e. $D\Delta u \delta W_{\text{ext}} (u, \delta u) = 0$), then the linearization of the virtual work is given by the linearization of the internal virtual work. For the current configuration, this results in [Holzapfel, 2001]

$$D\Delta u \delta W(u, \delta u) = \int_\Omega [\text{grad} \delta u : \text{grad} \Delta u \sigma + \text{grad} \delta u : Z : \text{grad} \Delta u] \, dv \quad \text{or}$$

$$= \int_\Omega \sum_{a,b,c,d=1}^3 \frac{\partial \delta u_a}{\partial x_b} (\delta_{ac}\sigma_{bd} + z_{abcd}) \frac{\partial \Delta u_c}{\partial x_d} \, dv.$$ \hspace{1cm} (3.58)

The term $\delta_{ac}\sigma_{bd} + z_{abcd}$ represents the effective elastic tensor, which has the nature of the tangent stiffness matrix. Its first term is the geometrical stress contribution and its second term is the material contribution.

Similarly, the linearization of the virtual work equation (3.28) for the reference configuration is [Holzapfel, 2001]

$$D\Delta u \delta W(u, \delta u) = \int_{\Omega_0} \left[ \text{Grad} \delta u : \text{Grad} \Delta u \mathbf{S} + \mathbf{F}^T \text{Grad} \delta u : Z : \mathbf{F}^T \text{Grad} \Delta u \right] \, dV \quad \text{or}$$

$$= \int_{\Omega_0} \sum_{A,B,C,D=1}^3 \frac{\partial \delta u_A}{\partial X_B} (\delta_{ac}S_{BD} + F_{aA}F_{cC}Z_{ABCD}) \frac{\partial \Delta u_c}{\partial X_D} \, dV,$$ \hspace{1cm} (3.59)

with effective elastic tensor $\delta_{ab}S_{BD} + F_{aA}F_{cC}Z_{ABCD}$.

Iterative approaches according to equations (3.58) and (3.59) are called updated Lagrangian and total Lagrangian, respectively. In the total Lagrangian approach, the reference configuration is kept equal to the initial configuration at time 0. In the updated Lagrangian
method, the final configuration of a load step becomes the reference configuration for the next load step. Identical results are obtained by both approaches [Bathe, 1996]. The non-linear models in this work were solved using the updated Lagrangian method, and only this method will be further described.

The linearized virtual work of the current configuration (3.58) can be expressed with respect to the nodal displacements via

\[
D_{\Delta \mathbf{u}} \delta W(\mathbf{u}, \delta \mathbf{u}) = \sum_{m=1}^{M} \int_{\Omega(m)} (\nabla \mathbf{H} \cdot \delta \mathbf{u} : \nabla \mathbf{H} \cdot \Delta \mathbf{u} \sigma + \nabla \mathbf{H} \cdot \delta \mathbf{u} : \varepsilon : \nabla \mathbf{H} \cdot \Delta \mathbf{u}) d\mathbf{v}^{(m)}
\]

or

\[
= \sum_{m=1}^{M} \int_{\Omega(m)} \sum_{a,c=1}^{3} \sum_{i=1}^{N} \frac{\partial H^i}{\partial x_b} \Delta \mathbf{u}_{ai} (\delta \mathbf{a}_c \sigma_{bd} + z_{abcd}) \sum_{j=1}^{N} \frac{\partial H^j}{\partial x_d} \Delta \mathbf{u}_{cj} d\mathbf{v}^{(m)}. \tag{3.60}
\]

Then \(D_{\Delta \mathbf{u}} \delta W(\mathbf{u}, \delta \mathbf{u})\) can be changed into matrix and vector notation \(\mathbf{K} \Delta \mathbf{u}\) by factoring out \(\delta \mathbf{u}_{ai}\) and \(\Delta \mathbf{u}_{cj}\), cancelling \(\delta \mathbf{u}_{ai}\), defining the element tangent stiffness tensor \(k^{(m)}_{iacj}\) as

\[
k^{(m)}_{iacj} = \int_{\Omega(m)} \frac{\partial H_i}{\partial x_b} (\delta \mathbf{a}_c \sigma_{bd} + z_{abcd}) \frac{\partial H_j}{\partial x_d} d\mathbf{v}^{(m)}, \tag{3.61}
\]

defining the system tangent stiffness tensor \(\mathbf{K}_{iacj} = \sum_{m=1}^{M} k^{(m)}_{iacj}\), and rearranging \(\mathbf{K}_{iacj}\) (similar to equations (3.52)-(3.56)). The linearized equation can then be written as a system of linear equations

\[
\begin{align*}
\mathbf{K} \Delta \mathbf{u} & = \mathbf{f} - \mathbf{K} \Delta \mathbf{u}, \\
\sum_{p=1}^{3N} \mathbf{K}_{pq} \Delta \mathbf{u}_p & = f_q - \sum_{p=1}^{3N} \mathbf{K}_{pq} \Delta \mathbf{u}_p & \text{for } x_p \text{ not on } \partial \Omega^u, \quad q = 1, 2, ..., 3N, \\
\Delta \mathbf{u}_p & = \mathbf{u}(x_p) & \text{for } x_p \text{ on } \partial \Omega^u,
\end{align*}
\]

with the right hand side as defined in equation (3.56).

The Newton-Raphson method can be employed to obtain from equation (3.62) an approximate solution for the static non-linear elastic problem. Let \(i\) be the subscript of the current equilibrium equation. The iterative approach can then be summarized by [ANSYS Inc., 2003]

\[
\mathbf{K}_i \Delta \mathbf{u}_i = \mathbf{f} - \mathbf{K}_i \Delta \mathbf{u}_i \tag{3.63}
\]

and

1. Assume initial nodal displacement field \(\Delta \mathbf{u}_0\) (zero at first time step, converged solution from previous time step), set \(i = 0\),

2. Compute updated tangent stiffness matrix \(\mathbf{K}_i\) and restoring load \(\mathbf{K}_i \Delta \mathbf{u}_i\) from \(\Delta \mathbf{u}_i\),
3. Calculate $\Delta \hat{u}_i$ from equation (3.63),

4. Set $\hat{u}_{i+1} = \hat{u}_i + \Delta \hat{u}_i$ to obtain next approximation of displacement field, set $i = i + 1$,

5. Repeat steps 2 to 4 until convergence (e.g. $\|f - K_i \hat{u}_i\| < \epsilon_R \|f\|$ or $\|\Delta \hat{u}_i\| < \epsilon_u \|\hat{u}\|$ where $\epsilon_R$ and $\epsilon_u$ are tolerances).

Depending on how often the tangent stiffness matrix $K_i$ is updated, several Newton-Raphson solution procedures exist. In the full Newton-Raphson procedure, $K_i$ is updated at every iteration. In the modified Newton-Raphson procedure, $K_i$ is only updated during the first iteration of each substep for a static analysis [ANSYS Inc., 2003].

The finite element formulation for static linear and non-linear elasticity problems described in this and the last section are based on a single type of solution variables - the displacements. While this is the most common approach, the usefulness of introducing another type of solution variable for nearly incompressible materials is examined next.

### 3.4.4 Volume Preservation

The pure displacement based method from Section 3.4.2 and 3.4.3 is computationally efficient. Yet, its accuracy depends on the level of volume preservation. For example, the constitutive relationship of isotropic elastic material can be described by $\sigma_{ab} = \kappa e_v \delta_{ab} + 2\mu e'_{ab}$, with bulk modulus $\kappa = E/[3(1-2v)]$, shear modulus $\mu = E/[2(1+v)]$, volumetric strain $e_v = \text{tr}(e)$ and deviatoric strain components $e'_{ab} = e_{ab} - 1/3 e_v \delta_{ab}$ for $a, b = 1, 2, 3$. As the material conserves more volume, Poisson’s ratio $v$ approaches 0.5, $\kappa$ increases to infinity and $e_v$ decreases to zero. Under nearly incompressible conditions, small errors in the predicted volumetric strain $e_v$ will therefore appear as large errors in the hydrostatic pressure $p = -\text{tr}(\sigma)/3$. This will also affect the displacement prediction, since external loads are balanced by the stresses. This may result in locking, where predicted displacements are much smaller than they should be, or no convergence at all. The situation could be improved by increasing the FEM mesh resolution, which generally leads to reduced prediction errors of the volumetric strain [Bathe, 1996]. Thus, displacement based FEMs could still provide accurate solutions of nearly incompressible materials by employing very fine meshes. Yet, fully incompressible deformations cannot be handled in this way, since any prediction error of the volumetric strain will lead to an infinite hydrostatic pressure. Instead, the principle of virtual displacements is extended by introducing another solution variable, the virtual pressure $\delta p$. Note that the accuracy of this so-called mixed u-p formulation is independent of the level of compressibility.

The virtual work equation for the current configuration (3.26) can be expressed in terms
of the independent variables $u$ and $p$ by

$$\int_\Omega \sum_{a,b=1}^3 \sigma'_{ab} \delta \epsilon'_{ab} \, dv - \int_\Omega p \delta e_v \, dv = \delta W_{\text{ext}}, \quad (3.64)$$

with deviatoric stress $\sigma'_{ab} = \sigma_{ab} + p \delta_{ab}$. Using $p = -\kappa e_v$, the two independent variables can be connected by the constraint [Bathe, 1996]

$$\int_\Omega \left( e_v + \frac{p}{\kappa} \right) \delta p \, dv = \int_\Omega e_v \delta p \, dv + \int_\Omega \frac{p}{\kappa} \delta p \, dv = 0. \quad (3.65)$$

Analogous to Section 3.4.2, the finite element formulation of equation (3.64) and (3.65) results in the system of linear equations [Bathe, 1996]

$$\begin{bmatrix} K_{uu} & K_{up} \\ K_{pu} & K_{pp} \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}, \quad (3.66)$$

where $\hat{u}$ and $\hat{p}$ denote the unknown nodal displacements and pressures, respectively,

$$K_{uu} = \sum_{m=1}^M \int_{\Omega^{(m)}} \text{grad} \mathbf{H}' : \mathbf{z}' : \text{grad} \mathbf{H}' \, dv^{(m)},$$

$$K_{up} = K_{pu}^T = -\sum_{m=1}^M \int_{\Omega^{(m)}} \text{grad} \mathbf{H}_v \cdot \mathbf{H}_p \, dv^{(m)}, \quad (3.67)$$

$$K_{pp} = -\sum_{m=1}^M \int_{\Omega^{(m)}} \frac{1}{\kappa} \mathbf{H}_p \cdot \mathbf{H}_p \, dv^{(m)},$$

$$f = \sum_{m=1}^M \int_{\Omega^{(m)}} \rho b \cdot \mathbf{H} \, dv^{(m)} + \sum_{m=1}^M \int_{\partial \Omega^{(m)}} \hat{t} \cdot \mathbf{H} \, ds^{(m)},$$

$$e' = \text{sym}(\text{grad} \mathbf{H}' \cdot \hat{u}), \quad e_v = \text{sym}(\text{grad} \mathbf{H}_v \cdot \hat{u}), \quad p = \mathbf{H}_p \hat{p}, \quad \text{and} \quad \mathbf{z}' \text{ represents the elastic tensor for the deviatoric stress and strain components.}$$

For fully incompressible materials ($v = 0.5$), $\kappa$ is infinite and (3.65) becomes $\int_\Omega e_v \delta p \, dv = 0$. Consequently, (3.66) turns into

$$\begin{bmatrix} K_{uu} & K_{up} \\ K_{pu} & 0 \end{bmatrix} \begin{bmatrix} \hat{u} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} f \\ 0 \end{bmatrix}. \quad (3.68)$$

Fewer unknown nodal pressures $\hat{p}$ than displacements $\hat{u}_a$ should be used to avoid a situation where the elements behave like displacement-based elements. Pressure is therefore defined per element or for a subset of element nodes.

### 3.4.5 Summary

This section discussed finite element formulations for static linear and non-linear elasticity problems as well as for incompressible materials. This concludes the general review of the main components of mechanical computer models, which also included descriptions on continuum mechanics (Section 3.2) and material models (Section 3.3). Breast specific material properties and mechanical models are discussed in the remainder.
3.5 Mechanical Breast Properties

This section reviews what is known about the mechanical properties of the breast. The anatomy of the breast was described in Section 1.2. The important structures of the breast from the mechanical point of view are the fatty tissue, ductal and glandular tissue, connective tissue, Cooper’s ligaments, skin and if applicable tumourous tissue, see Figure 1.1 on page 26.

The composition of the breast, and hence its mechanical characteristics, changes under hormonal influence. Ductal and glandular epithelial cells increase in number and in size in the premenstrual phase, during pregnancy and with hormone replacement therapy (HRT)\(^4\). The opposite happens in the postmenstrual phase, with the cessation of nursing, with the commence of the menopause and with the termination of HRT. These effects will manifest themselves on MR images as an alteration of fibroglandular tissue volume [Graham et al., 1995], and hence can be accounted for. Other elasticity changes will not lead to noticeable MR intensity changes. These include stretching or softening of Cooper’s ligaments, which may occur during pregnancy or with old age, and replacement of connective tissue by ductal or glandular tissue.

The elastic values for the individual tissue types are typically derived from ex-vivo experiments on small specimens. Only a few studies have been reported for the breast [Sarvazyan et al., 1994; Krouskop et al., 1998; Wellman, 1999]. Their measurements were based on compressing thin sliced specimens with an indenter of 4 to 5mm diameter. A relatively small indenter was used to avoid heterogeneous tissue and an irregular surface. The response to other load scenarios was not tested. Specimens are unlikely to consist of purely ductal, glandular or connective tissue. Elastic properties of fibroglandular tissue will therefore vary depending on the proportion of the individual components. The mechanical properties of the tissue are likely to change after excision because the sample is not in its natural physiological state. For example, the blood supply and the body fluids are absent, the fibrous support system is possibly damaged and the tissue is normally tested at room temperature (21-24°C) instead of body temperature. Further information may come from the field of elastography, where researchers try to deduce the elastic property distribution from the in-vivo measured displacement field of the tissue for known applied forces to the whole organ [Lawrence et al., 1998a; Lawrence et al., 1999; Sinkus et al., 2000; McKnight et al., 2001; van Houten et al., 2003].

\(^4\)HRT describes the use of female hormones for the relief of symptoms resulting from cessation of ovarian function, e.g. at the time of the menopause.
Table 3.2 summarizes the reported Young’s moduli for fatty, fibroglandular and tumourous tissue. The reported tumourous tissue was mainly invasive ductal carcinoma. These values vary greatly. In all studies, Young’s moduli increased from fatty to fibroglandular tissue and from fibroglandular to tumourous tissue. Note that there appear to be inconsistencies in [Wellman, 1999] regarding the reported Young’s moduli, the shown graphs and the fitted exponential curves.

[Krouskop et al., 1998] and [Wellman, 1999] assessed breast specimens for different pre-staining states and load frequencies. They observed that all tissue types apart from fat had clearly non-linear stress-strain relationships. Of these, tumourous tissue displayed the greatest increase of Young’s moduli for higher strains. Breast tissue displayed primarily an elastic response, with changes in indentation velocity altering Young’s moduli by less than 5% [Krouskop et al., 1998; Wellman, 1999]. Cancer specimens became much softer once strained by more than 5% [Krouskop et al., 1998] or 10% [Wellman, 1999]. A follow-up study confirmed this behaviour for infiltrating ductal carcinomas without lobular features [Krouskop et al., 2003]. The sensitivity of Young’s modulus to a temperature change from 24°C to 38°C was investigated in [Krouskop et al., 1998]. They found that the viscous behaviour of 4 fatty specimens changed, while Young’s moduli remained within the range of measured values. No changes of Young’s moduli were measurable for 2 non-fatty specimens. Results remained unchanged when tests were repeated after 90 minutes [Krouskop et al., 1998]. The recorded Young’s moduli for a tissue type were more consistent for samples from the same breast than for samples from different breasts [Sarvazyan et al., 1994]. Methods for estimating the parameters of hyperelastic models for ex-vivo experiments were proposed in recent publications [Samani and Plewes, 2004; Kerdok et al., 2005]. Parameters were optimized such that the correspondence between a non-linear FEM model and the experiment was maximized. Application of these methods to a large number of breast samples will hopefully clarify the currently conflicting results.

Initial results from the field of elastography vary substantially. Most methods were 2D and hence limited in accuracy due to the negligence of through plane motion. Breast tissue was commonly assumed to be isotropic and linear elastic, which is a known simplification. Breasts were often compressed by an uncontrolled amount, leading to variations in the pre-staining and hence in Young’s modulus. A range of loading conditions and reconstruction techniques were employed, and relative few cases were studied. Incompressibility assumptions could have been violated if forces were applied that exceeded venous blood pressure values. While elastography can certainly provide information for detecting stiffer regions in a breast,
<table>
<thead>
<tr>
<th>Reference</th>
<th>Number&lt;sup&gt;a&lt;/sup&gt; of cases</th>
<th>Number of fatty/fibroglandular/tumourous specimens; in-vivo: number of volunteers (v) or patients (p)</th>
<th>Frq (Hz) (μm) (%)</th>
<th>Dsp</th>
<th>e&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Young’s modulus in kPa (mean±SD)</th>
<th>E&lt;sub&gt;f&lt;/sub&gt;</th>
<th>E&lt;sub&gt;g&lt;/sub&gt;</th>
<th>E&lt;sub&gt;t&lt;/sub&gt;</th>
<th>E&lt;sub&gt;f&lt;/sub&gt;/E&lt;sub&gt;f&lt;/sub&gt;</th>
<th>E&lt;sub&gt;g&lt;/sub&gt;/E&lt;sub&gt;f&lt;/sub&gt;</th>
<th>Comment&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Sarvazyan et al., 1994]</td>
<td>56/42/18</td>
<td>Ex-vivo: number of fatty/fibroglandular/tumourous specimens; in-vivo: number of volunteers (v) or patients (p)</td>
<td>0.8 ± 0.3</td>
<td>1.2 ± 0.3</td>
<td>3.6 ± 0.6</td>
<td>1.5</td>
<td>4.5</td>
<td>1.6</td>
<td>2.4</td>
<td>3.6</td>
<td>9.8</td>
<td>DCIS</td>
</tr>
<tr>
<td>[Krouskop et al., 1998]</td>
<td>8/31/32</td>
<td>Frq: frequency, Dsp: maximum induced displacement, e: amount of pre-straining</td>
<td>18.0 ± 7.0</td>
<td>28.0 ± 14.0</td>
<td>106.0 ± 32.0</td>
<td>1.6</td>
<td>5.9</td>
<td>1.6</td>
<td>2.4</td>
<td>3.6</td>
<td>9.8</td>
<td>IDC</td>
</tr>
<tr>
<td>[Krouskop et al., 1998]</td>
<td>8/31/32</td>
<td></td>
<td>20.0 ± 8.0</td>
<td>48.0 ± 15.0</td>
<td>558.0 ± 180.0</td>
<td>2.4</td>
<td>27.9</td>
<td>5.0</td>
<td>17.5</td>
<td>9.8</td>
<td>17.5</td>
<td>IDC</td>
</tr>
<tr>
<td>[Wellman, 1999]</td>
<td>26/7/25</td>
<td></td>
<td>4.8 ± 2.5</td>
<td>17.5 ± 8.6</td>
<td>47.1 ± 19.8</td>
<td>3.6</td>
<td>9.8</td>
<td>3.6</td>
<td>27.9</td>
<td>9.8</td>
<td>17.5</td>
<td>IDC</td>
</tr>
<tr>
<td>[Wellman, 1999]</td>
<td>26/7/25</td>
<td></td>
<td>6.6 ± 3.7</td>
<td>33.3 ± 12.0</td>
<td>115.7 ± 42.9</td>
<td>5.0</td>
<td>17.5</td>
<td>5.0</td>
<td>17.5</td>
<td>9.8</td>
<td>17.5</td>
<td>IDC</td>
</tr>
<tr>
<td>[Wellman, 1999]</td>
<td>26/7/25</td>
<td></td>
<td>10.4 ± 7.9</td>
<td>88.1 ± 66.7</td>
<td>384.5 ± 126.9</td>
<td>8.5</td>
<td>37.0</td>
<td>8.5</td>
<td>37.0</td>
<td>9.8</td>
<td>17.5</td>
<td>IDC</td>
</tr>
<tr>
<td>[Wellman, 1999]</td>
<td>26/7/25</td>
<td></td>
<td>17.4 ± 8.4</td>
<td>271.8 ± 167.7</td>
<td>1366.5 ± 348.2</td>
<td>15.6</td>
<td>78.5</td>
<td>15.6</td>
<td>78.5</td>
<td>9.8</td>
<td>17.5</td>
<td>IDC</td>
</tr>
<tr>
<td>[Lawrence et al., 1998a]</td>
<td>9v</td>
<td></td>
<td>1.3 ± 0.2</td>
<td>7.4 ± 0.6</td>
<td>5.7</td>
<td>from 2D shear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>from 2D shear</td>
</tr>
<tr>
<td>[Lawrence et al., 1999]</td>
<td>7p</td>
<td></td>
<td>7.5 ± 3.0</td>
<td>18.0 ± 4.5</td>
<td>75.0 ± 54.0</td>
<td>2.4</td>
<td>10.0</td>
<td>2.4</td>
<td>10.0</td>
<td>9.8</td>
<td>17.5</td>
<td>from 2D shear</td>
</tr>
<tr>
<td>[Sinkus et al., 2000]</td>
<td>1p</td>
<td></td>
<td>0.7</td>
<td>2.1</td>
<td>3.0</td>
<td>3D, anisotropic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3D, anisotropic</td>
</tr>
<tr>
<td>[Kruse et al., 2000]</td>
<td>1p</td>
<td></td>
<td>20.0</td>
<td>37.5</td>
<td>62.5</td>
<td>1.9</td>
<td>3.1</td>
<td>1.9</td>
<td>3.1</td>
<td>9.8</td>
<td>17.5</td>
<td>from 2D shear</td>
</tr>
<tr>
<td>[McKnight et al., 2001]</td>
<td>6v</td>
<td></td>
<td>9.9 ± 5.7</td>
<td>22.5 ± 10.8</td>
<td>2.3</td>
<td>from 2D shear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>from 2D shear</td>
</tr>
<tr>
<td>[McKnight et al., 2001]</td>
<td>6p</td>
<td></td>
<td>24.0 ± 12.0</td>
<td>99.0 ± 87.0</td>
<td>4.1</td>
<td>from 2D shear</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>from 2D shear</td>
</tr>
<tr>
<td>[van Houten et al., 2003]</td>
<td>5v</td>
<td></td>
<td>20.9 ± 2.7</td>
<td>26.3 ± 2.3</td>
<td>1.3</td>
<td>3D</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3D</td>
</tr>
<tr>
<td>[Sinkus et al., 2005]</td>
<td>1p</td>
<td></td>
<td>3.3</td>
<td>17.4</td>
<td>5.3</td>
<td>3D, anisotropic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3D, anisotropic</td>
</tr>
</tbody>
</table>

Table 3.2: Overview of reported Young’s moduli for fatty (E<sub>f</sub>), fibroglandular (E<sub>g</sub>) and tumourous (E<sub>t</sub>) tissue.

<sup>a</sup>Ex-vivo: number of fatty/fibroglandular/tumourous specimens; in-vivo: number of volunteers (v) or patients (p)

<sup>b</sup>Frq: frequency, Dsp: maximum induced displacement, e: amount of pre-straining

<sup>c</sup>DCIS: ductal carcinoma in situ, IDC: infiltrating ductal carcinoma; from shear: assuming incompressible, linear material, shear modulus μ was converted into Young’s modulus E by E = 3μ

<sup>d</sup>This Young’s modulus was extrapolated, because tumour specimens were damaged at strains higher than 10%.
it has yet to show its ability of determining accurate elasticity properties of the breast.

In conclusion, Young’s modulus increased from fatty to fibroglandular tissue and from fibroglandular tissue to tumourous tissue in all studies. The relative changes as well as the absolute measures were however inconsistent. Factors contributing to this situation include complexity of breast composition, invalid assumptions, variation in experimental setup and choice of estimation technique. Despite the uncertainty regarding the elastic properties of the breast, mechanical computer models of the breast have been developed. These are reviewed in the next section.

3.6 Review of Biomechanical Breast Models

Mechanical computer models are employed in the field of medical image processing for a variety of applications, including image guided interventions [Platenik et al., 2002], surgical simulation [Liu et al., 2003], treatment planning [Chabanas et al., 2003] and image registration [Hagemann et al., 1999; Ferrant et al., 2001]. The interaction between images and mechanical models are generally complementary. Mechanical models can supplement limited image information for predicting deformations or for constraining registrations, while images captured during the motion of an organ can be employed for building and verifying mechanical models.

In the case of the breast, biomechanical computer models employing FEMs have been explored for predicting mechanical deformations during a biopsy procedure [Azar et al., 2002], for modeling compressions similar to X-ray mammography [Samani et al., 2001b; Yin et al., 2004; Pathmanathan et al., 2004], for the registration of MR and X-ray mammograms [Ruitter, 2003], for testing reconstruction algorithms in elastography [Sinkus et al., 2000; van Houten et al., 2001; Samani et al., 2001a], and as a forward model for elastography [Washington and Miga, 2004]. The different biomechanical breast models vary mainly with respect to the mesh generation, the boundary conditions employed, the assumed tissue properties and the solution strategies.

3.6.1 Mesh Generation

Table 3.3 lists the mesh sizes and element types used by biomechanical breast models. The meshes had between 343 and 14902 hexahedra. [Yin et al., 2004] generated elements directly from blocks of $6 \times 6 \times 5$ voxels, which resulted in a very irregular surface with many right angles. Smoother surfaces were achieved by mapping a mesh of hexahedra to the approximated continuous breast boundary in all other cases. All elements had linear shape functions.
### 3.6 Review of Biomechanical Breast Models

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of elements</th>
<th>Volume elements</th>
<th>Surface elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Azar et al., 2002]</td>
<td>$7 \times 7 \times 57 = 2793$</td>
<td>8-noded hexahedra</td>
<td>3-noded triangles</td>
</tr>
<tr>
<td>[Samani et al., 2001b]</td>
<td>$10 \times 12 \times 19 = 2280$</td>
<td>8-noded hexahedra</td>
<td>4-noded quadrilaterals</td>
</tr>
<tr>
<td>[Ruiter, 2003]</td>
<td>$7 \times 7 \times 7 = 343$</td>
<td>8-noded hexahedra</td>
<td>4-noded quadrilaterals</td>
</tr>
<tr>
<td>[Pathmanathan et al., 2004]</td>
<td>$12 \times 12 \times 5 = 4608$</td>
<td>8-noded hexahedra</td>
<td>-</td>
</tr>
<tr>
<td>[Yin et al., 2004]</td>
<td>14902</td>
<td>8-noded hexahedra</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3.3: Mesh configurations of reported biomechanical breast models.

#### 3.6.2 Boundary Conditions

Most models employed only displacement boundary conditions, while [Pathmanathan et al., 2004] simulated the response to gravity. Nodes adjacent to the pectoral muscle were constrained to have zero displacements in all models. Breast compression was modelled as a contact problem [Samani et al., 2001b; Yin et al., 2004], by prescribed surface displacements [Azar et al., 2002] or by employing a penalty function [Pathmanathan et al., 2004]. Fixed plates were assumed to have zero displacements [Azar et al., 2002; Samani et al., 2001b]. [Azar et al., 2002] assumed that the mobile plates align with the sagittal-plane and move perpendicular, which may not hold exactly. [Ruiter, 2003] compared various boundary conditions, including surface displacements from image registration, one- and two-sided plate compression formulated as a contact problem or directly by displacements, and combined strategies.

#### 3.6.3 Tissue Properties

Biomechanical breast models generally accounted for fatty, fibroglandular and skin tissue, and sometimes also for tumourous tissue [Yin et al., 2004]. All tissue types were modelled as homogeneous, isotropic materials. Figure 3.9 depicts the reported stress-strain relationships and the ex-vivo data they were based on. The reported Young’s moduli or strain energy functions are listed in Table 3.4. [Samani et al., 2001b] and [Yin et al., 2004] modelled fat similar to the measured ex-vivo results. [Azar et al., 2002] compensated for the effects of Cooper’s ligaments by increasing the stiffness of the fatty tissue. This function was initially of a different form [Azar et al., 2000]. [Pathmanathan et al., 2004] modelled all materials as fibroglandular tissue. The ex-vivo results for fibroglandular tissue were quite different at higher strains [Wellman, 1999; Krouskop et al., 1998]. The exponential fit provided by [Wellman, 1999] does not seem to approximate the measured data well. The neo-Hookean model is too stiff for low strains [Samani et al., 2001b]. [Ruiter, 2003] compared the performance of several of these material models. If necessary, she fitted linear, neo-Hookean and exponential...
functions to match the constitutional relationship at 21% strain.

If only displacement boundary conditions are applied, then models predict the same deformation as long as the tissue types have the same ratios of Young’s moduli. Only the stress values are affected by the actual Young’s moduli in this case. Figure 3.10 depicts the Young’s moduli ratios for strains up to 25%. Four models assumed that the fibroglandular tissue is quite similar to fatty tissue, with average ratios ranging from 0.84 to 1.75, see Table 3.4.
### 3.6 Review of Biomechanical Breast Models

Young’s modulus $E$ or strain energy function $\Psi$ in kPa for

<table>
<thead>
<tr>
<th>Reference</th>
<th>Young’s modulus $E$ or strain energy function $\Psi$ in kPa for</th>
<th>$\frac{E_g}{E_f}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fatty tissue</td>
<td>Fibro glandular tissue</td>
</tr>
<tr>
<td>[Azar et al., 2000]$^1$</td>
<td>$E_f = 718.16e + 4.46$ if $e &lt; 0.25$</td>
<td>$E_g = 15.1 \exp(10.0e)$</td>
</tr>
<tr>
<td></td>
<td>$E_f = 184$ if $e \geq 0.25$</td>
<td></td>
</tr>
<tr>
<td>[Azar et al., 2002]$^1$</td>
<td>$E_f = 1702.1e^2 + 166.11e + 4.46$ if $e &lt; 0.15$</td>
<td>$E_g = 15.1 \exp(10.0e)$</td>
</tr>
<tr>
<td></td>
<td>$E_f = 15.1 \exp(10.0e)$ if $e \geq 0.15$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Samani et al., 2001b]$^1$</td>
<td>$E_f = 519.7e^2 + 2.4e + 4.9$</td>
<td>$E_g = 123888.9e^3 - 11766.7e^2 + 696.9e + 12.1$</td>
</tr>
<tr>
<td>neo-Hookean model$^a$</td>
<td>$\Psi_f = 3.57(I_1 - 3)$</td>
<td>$\Psi_g = 105.10(I_1 - 3)$</td>
</tr>
<tr>
<td>[Pathmanathan et al., 2004]$^1$</td>
<td>$\Psi_f = \Psi_g$</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$\Psi_f = 2(I_1 - 3) + 1.333(I_2 - 3)$</td>
<td></td>
</tr>
<tr>
<td>[Yin et al., 2004]$^2$</td>
<td>$\Psi_f = 2(I_1 - 3) + 1.333(I_2 - 3)$</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.4: Young’s modulus or strain energy function of fat, fibro glandular tissue and skin for biomechanical breast models from the literature adapting or approximating ex-vivo results from $^1$[Wellman, 1999] and $^2$[Krouskop et al., 1998]. All models assumed isotropic materials. The last column states mean($E_g/E_f$) for $e \in [0, 0.25]$.

$^a$The parameters were derived from fitting a neo-Hookean model to the stated polynomial using [ANSYS Inc., 2000], since no values were given in [Samani et al., 2001b].
3.6 Review of Biomechanical Breast Models

The ex-vivo measurements from [Wellman, 1999] show a substantial increase in relative stiffness with increased strains. This is best modelled by the polynomials from [Samani et al., 2001b] and less well approximated by the exponential functions defined in [Wellman, 1999]. Neo-Hookean models cannot describe this effect.

Tissue incompressibility was supported by the use of hyperelastic models [Samani et al., 2001b; Ruiter, 2003; Pathmanathan et al., 2004; Yin et al., 2004] or by Poisson’s ratios \( v \) close to 0.5 \( (v = 0.49999 \text{ [Azar et al., 2002], } v \in [0.45, 0.499] \text{ [Ruiter, 2003]} \).

3.6.4 Solution Strategies

[Samani et al., 2001b; Pathmanathan et al., 2004; Ruiter, 2003] employed a finite deformation formulation to account for large breast deformations. The updated Lagrangian approach was solved iteratively using Newton’s method, see Section 3.4.3. [Azar et al., 2002] and [Yin et al., 2004] divided the large deformation process into smaller incremental displacement steps of 2.6% and 1% strain, respectively, and used the small strain formulation for every step. This approach is fast, but neglects the geometrical stress contribution of the individual steps. [Ruiter, 2003; Pathmanathan et al., 2004] employed a mixed u-p formulation (see Section 3.4.4) to improve convergence and accuracy for (nearly) incompressible materials.

3.6.5 Evaluation

Validation of biomechanical breast models is a difficult task since the true displacements are unknown. So far, models have been assessed based on the predicted location of anatomical landmarks selected in breast images acquired before and after in-vivo compression [Azar et al., 2001; Azar et al., 2002; Ruiter, 2003], by visual comparison of the simulated compressed breast image with the uncompressed breast [Samani et al., 2001b], or not at all [Yin et al., 2004; Pathmanathan et al., 2004].

[Azar et al., 2000] reported an average (maximum) displacement error of 4.3mm (5.0mm) for predicting the location of 3 landmarks which were initially displaced by 8.3mm (10.1mm). Changing the material properties of fatty tissue and adding skin resulted in an average (maximum) error to 3.3mm (4.6mm) [Azar et al., 2002]. Applying this technique to 3 patients, the locations of their lesions were predicted with an average (maximum) error of 3.7mm (7.3mm) [Azar et al., 2001]. The average distance of 16 landmarks was reduced from 18.3mm to 2.5mm for the best configuration from [Ruiter, 2003]. The maximum error decreased from 25.7mm to 4.0mm for this setup. The better performance by [Ruiter, 2003] was mainly due to improved boundary conditions. The assessment was either based on very few landmarks or only on a single breast. The observer variability of selecting landmarks was not evaluated.
3.7 Summary

Continuum mechanics, material models, FEMs, mechanical breast properties and biomechanical breast models were reviewed in this chapter. The review showed that large deformations require a finite strain formulation and that incompressible tissue is best modelled by a mixed u-p formulation and hyperelastic material models. Transverse isotropic materials, which have not been employed previously for the breast, are interesting for modelling simple anisotropic behaviour.

All studies which investigated the mechanical breast properties reported an increased stiffness from fatty to fibroglandular tissue and from fibroglandular tissue to tumourous tissue. Yet, the estimated elastic values were inconsistent and only uniaxial experiments were conducted. Two ex-vivo studies showed that fibroglandular and tumourous tissue have non-linear constitutive laws. Some variations of elastographic estimates can therefore be explained by differences in pre-straining.

Biomechanical breast models have been built for several applications. Only a single material model and solution strategy was considered by these studies apart from [Ruiter, 2003], whose model assessment was inspired by the author’s work [Tanner et al., 2001; Tanner et al., 2002a]. The evaluation of the model performance was otherwise very limited.

The inconsistency in the reported breast material properties precludes that a model configuration can be chosen based on these values. Furthermore, the influence of other aspects like the mesh resolution, the boundary conditions and the solution strategy on the accuracy of these models have generally not been evaluated. It is therefore necessary to assess the accuracy of the employed model for a range of configurations before simulating plausible breast deformations for validating the registration of DCE MR mammograms. This assessment and the design of the models are described in Chapter 5.
Chapter 4

Review of MR Breast Lesion Segmentation and Classification

This chapter reviews automatic methods for segmenting and classifying MR breast lesions. The aim is to find a suitable framework for the computer aided diagnosis system described in Chapter 7.

4.1 Introduction

Interpreting dynamic contrast-enhanced MR mammograms for signs of breast cancer is time consuming and complex since the whole image sequence needs to be analysed in detail. Misclassifications can arise from either overlooking a suspicious region or from incorrectly interpreting a suspicious region. In recent studies, radiologists achieved on average a high sensitivity (93%) but a moderate specificity (68%) when reading DCE MR mammograms, see Table 1.2. The high sensitivity indicates that the detection of suspicious regions is the easier task. This thesis is therefore focused on supporting the radiologist in the classification of MR breast lesions by means of image processing technologies.

Figure 4.1 shows the processing steps of the proposed computer aided diagnosis (CAD) system for dynamic contrast-enhanced MR breast images. All tasks are fully automated apart from the crude manual outlining, which can also be based on the unregistered images if necessary. In comparison to other CAD systems, image registration is included, since classification is based on the properties of an image sequence which could have been affected by patient motion.

The image registration performance is assessed and a suitable configuration is identified.

Figure 4.1: Computer aided diagnosis system for dynamic contrast-enhanced MR mammography.
in Chapter 6. This chapter reviews segmentation methods (Section 4.2); feature candidates and feature selection methods (Section 4.3); and classification approaches (Section 4.4) employed in this context. Chapter 7 describes the proposed CAD system and its classification performance.

### 4.2 Segmentation

This section describes the problem of segmenting lesions from DCE MR mammograms, provides a short overview of image segmentation methods and reviews segmentation algorithms proposed for DCE MR mammography. General reviews on medical image segmentation, deformable models and clustering can be found in [Pham et al., 2000; McInerney and Terzopoulos, 1996; Montagnat et al., 2001; Jain et al., 1999].

#### 4.2.1 Introduction

The development of automatic CAD systems for DCE MR mammography relies on the collection of ground truth information of the breast lesion’s image position and extent. Currently, a radiologist’s segmentation is the accepted gold standard for this definition. However, manual segmentation is prone to inaccuracies and is laborious. The required work and the present shortage of radiologists hamper the collection of enough segmentation examples such that reliable CAD systems can be created.

While humans can rapidly perceive image objects, the exact definition of the object boundary is time consuming, especially for contrast-enhanced image sequences. Fully automatic segmentation, on the other hand, has proven more difficult than expected. The aim in this thesis is therefore to develop a semi-automatic method that reduces the segmentation workload significantly, which is also applicable for weakly enhanced structures and which can readily be applied to unregistered and registered images without requiring additional manual segmentations.

#### 4.2.2 Overview of Segmentation Methods

Image segmentation is defined as the partitioning of an image into its constituent regions or objects. When the constraint that the regions are connected is removed, then segmentation is called pixel classification and the partitions are called classes. Segmentation methods can generally be categorized into discontinuity-based and similarity-based approaches [Gonzalez and Woods, 1993].

Discontinuity-based methods concentrate on the sharp local changes in image intensity. Methods include edge detection followed by edge linking and the watershed algorithm, which
uses edge detection and mathematical morphology to divide the image into homogeneous regions. Discontinuity-based approaches suffer from false and missing edges since intensity gradients are more affected by noise than the image intensity. Deformable models, like snakes, balloons and level-sets, have been developed to overcome this problem. They are based on deforming a closed boundary under the influence of shape-based forces and image-derived forces [McInerney and Terzopoulos, 1996; Montagnat \textit{et al.}, 2001]. The shape-based force can incorporate a priori knowledge about the location, size and shape of the structure. Various shape models have been proposed ranging from general smoothness cost functions to application specific statistical shape templates learned from a training set.

Similarity-based methods rely on the homogeneity characteristics of a set of pixels. Using the global intensity statistics, pixel can be unsupervised classified by for example simple thresholding, K-means clustering or the EM algorithm [Pham \textit{et al.}, 2000]. Difficulties arise from noise, image inhomogeneities, unknown number of clusters, sensitivity of the cluster algorithms to the initialization and wrong assumptions regarding the data distribution. Improvements can be achieved by incorporating prior knowledge from manually segmented images. Examples of such supervised methods are K-nearest-neighbour classifiers, Parzen window and Bayes classifiers. Pixel classifiers do generally not perform any spatial modelling and hence are sensitive to noise. Markov random field models, which are statistical models that define the relationship between nearby pixels, are therefore often incorporated to improve the robustness to noise. While pixel classification relies on the global intensity statistics, similarity-based segmentation can be based on the statistics from the individual regions. In region growing, for example, a region may be expanded by neighbouring pixels if their intensity values lie within a certain standard deviation of the mean intensity value of the region. Drawbacks of region growing are wrongly connected regions due to partial-volume effects and noisy boundaries. Region competition [Zhu and Yuille, 1996], which combines deformable contours and region growing, has been proposed to overcome these problems. Both methods require good initial regions.

In summary, segmentations and pixel classifications simply based on image features generally fail due to noise, inhomogeneities and partial volume effects. Incorporation of prior knowledge in the form of shape, class or neighbourhood models have been successful for certain applications. In the case of segmenting DCE MR mammograms, shape and general class models are unlikely to be suitable due to the high variability of the lesions’ shape, size, location and image properties. Case specific class models, as derived for example from coarse manual segmentations, are a promising semi-automatic framework for more accurate
4.2 Segmentation

Similarity-based methods can relatively easily be extended to multichannel images, i.e., images with more than one measurement, by employing multivariate statistics. The proposed segmentation is therefore based on a semi-automatic similarity-based pixel classification with subsequent region extraction.

4.2.3 Review of DCE MR Mammography Segmentation

A few automatic and semi-automatic segmentation algorithms have been proposed for the extraction of breast lesions from dynamic contrast-enhanced MR images. Their characteristics are as follows:

- [Lucas-Quesada et al., 1996] recommended segmentation of MR breast lesions by manually thresholding the similarity map, which was generated from the normalized cross-correlation between the time-intensity plot of each voxel and a reference plot derived from a small user-defined region of interest. This temporal correlation method compared favourably to a multispectral analysis method. The multispectral analysis method used 20 to 30 manually selected lesion voxels to generate a lesion cluster in the 2D pre- to post-contrast intensities space by means of the K-nearest-neighbour algorithm. A sequential contouring algorithm was used in both cases to extract a lesion outline. Note that the multispectral analysis method may have been disadvantaged by exploiting only limited data from the temporal domain, while the temporal correlation method was dependent on a single reference enhancement curve.

- [Hayton, 1998] employed a multi-stage approach for detecting regions of focal enhancement on 2D slices within the breast region \( B \). For each pixel, the enhancement profile was first characterized by the enhancement norm defined as

\[
E(x) = \left( \frac{I_{\text{max}}(x)}{I_0(x)} - 1 \right)^2 + \left( \frac{T - t_{63\%}(x)}{T} \right)^2 + \lambda (I_{\text{max}}(x) - I_0(x)),
\]

(4.1)

where \( I_{\text{max}} \) represents the maximum image intensity, \( I_0 \) denotes the pre-contrast image intensity, \( t_{63\%} \) stands for the time taken for the enhancement to reach 0.63(\( I_{\text{max}}/I_0 - 1 \)), time \( T \) represents the duration of the MR scan and \( \lambda \) denotes a weighting factor. Secondly, a strong membrane function \( (M_S) \) and a weak membrane function \( (M_W) \) were fitted to the enhancement norm image employing the Graduated Non-Convexity algorithm of [Blake and Zisserman, 1987]. Thirdly, all pixels which satisfied

\[
M_W(x) - M_S(x) < 3\sigma_E, \quad \text{where } \sigma_E^2 = \text{mean}_{x \in B} (E(x) - M_W(x))^2,
\]

(4.2)

were marked as background and a sparse membrane \( (M_{SP}) \) was fitted to these background pixels. Fourthly, pixels which satisfied \( M_W(x) - M_{SP}(x) > 3\sigma_E \) were identified...
as significantly enhancing pixels (foreground). The foreground pixels were then categorized to belong to either a blob or a linear shape structure employing the primitive identification algorithm from [Gilles, 1998]. Thereafter, lesions were extracted by applying a region growing algorithm to the pixels from the blob category. Finally lesions whose maximum enhancement \( I_{\text{max}} / I_0 \) was less than 0.4 for all pixels or whose contrast \((E(\mathbf{x}) - M_{SP}(\mathbf{x}))/ (E(\mathbf{x}) + M_{SP}(\mathbf{x}))\) was less than 0.5 for all pixels were discarded. The lesions detected by this segmentation method for 41 patients were visually identified by a radiologist as either false enhancement, true enhancement or outside the breast. Lesions outside the breast were excluded from the assessment. Without registration, 62% of 567 lesions were true positives and 38% were false negatives. After registration, this improved to 89% true positives and 11% false negatives for 420 lesions. Diffuse or weak enhancing lesions may not be detected by this method, as it is aimed to detect focal enhancements.

- [Fischer et al., 1999a] clustered the intensity enhancement profiles employing self-organizing Kohonen maps. The cluster results were shown to the user for interrogation of the dynamic sequences. No segmentation was attempted.

- [Gilhuijs et al., 2002; Chen et al., 2004] reduced the image sequence to a single image by voxelwise computing the variance of the intensity values over time for all voxels within the breast. A spherical region of interest, which included a high variance region, was then automatically selected from spheres expanding outwards from a manually selected seed point in the suspicious lesion. Thereafter, a segmentation threshold was determined by maximizing the interclass variance between the high variance voxels and the background voxels within this spherical region of interest. A 3D segmentation was finally extracted by 6-point-connected volume growing. The segmentation result was visually verified and the location of the seed point was adjusted if necessary. The accuracy of this method is limited because the segmentation does not take account of the temporal sequence of the intensity changes. For example, slow enhancements can have a similar variance as fast enhancements with washout and hence cannot be discriminated. This method is otherwise difficult to comment on since no further details were provided in [Gilhuijs et al., 2002; Chen et al., 2004].

- [Jacobs et al., 2003] employed a K-means related clustering algorithm for extracting 4D feature vectors of fatty, fibroglandular and lesion tissue from T1- and T2-weighted images and 3D fat-suppressed T1-weighted pre- and post-contrast images. Lesion clas-
4.2 Segmentation

...fication was based on the angular separation from the fat feature vector. Fat and fibroglandular reference feature vectors were provided by the user. Extraction of lesion outlines was not attempted.

- [Ketsetzis and Brady, 2004] employed a hidden Markov random field to segment T1 parametric maps of 2D MR breast images. The images were firstly corrected for the bias field [Styner et al., 2000]. The background was then removed using Otsu’s thresholding method. The T1 values of the tissue classes were assumed to follow a multivariate Gaussian distribution function. A Markov random field based on a logistic model and a second order neighbouring system was employed to determine the prior class probabilities. The foreground was finally segmented into 8 tissue classes using initially a maximum a posteriori approach and then simulated annealing for iteratively refining the solution. Example results were shown for 3 patients. The segmented tissue classes were reported to be subsets of fatty and fibroglandular tissue. Segmentation of abnormalities was not attempted.

- [Xiaohua et al., 2005] proposed a combined segmentation and registration approach. Breast tissue was firstly segmented into fatty and non-fatty tissue based on a K-means clustering of the pre-contrast image. For a given transformation, the non-fatty tissue was then segmented into normal, benign and malignant tissue based on two kinetic features derived from fitting the bi-exponential pharmacokinetic model, defined in equation (2.38) on page 70, to the intensity change over time. The two kinetic features were the relative enhancement at 1 minute \(K_{rel1} = I_1/I_0 - 1\) and the washout rate at half-time \(K_{rhwo} = (I_6 - I_3)/3\). Segmentation was based on a maximum a posteriori approach, assuming independent Gaussian distributions of the kinetic features for the class conditional density function and a Markov random field model for the prior class probabilities. Class assignment for the class conditional density function was based on thresholding the kinetic features, namely malignant if \(K_{rel1} > 0.5\) and \(K_{rhwo} < 0\), benign if \(K_{rel1} > 0.3\) and \(K_{rhwo} < 0\) and normal if \(K_{rel1} < 0.3\) and \(|K_{rhwo}| < 0.05\). The combined segmentation and registration approach was applied to 2D images of 15 patients. Example slices were presented showing the reduction in the residual fitting error of the pharmacokinetic model and the change in tissue segmentation. Extraction of lesion outlines was not attempted.
4.2.4 Summary

In summary, only a few methods have been proposed for segmenting lesions from DCE MR mammograms. Often, these were based on clustering the time-intensity profiles without regard of the spatial relationship between voxels of the same cluster or on reducing the time-intensity profiles to one or two predefined features. Class conditional probabilities were generally assumed to be Gaussian distributed. In contrast, the segmentation refinement method proposed in this work, and described in Chapter 7, aims to extract the most probable lesion object of the contrast-enhanced MR image sequence from the data provided by a coarse manual segmentation and prior knowledge about the segmentation process, without assuming separability based on predefined kinetic features or a particular class conditional probability density function.

The next task after a lesion has been segmented is to extract features describing properties of the lesion, which might help in discriminating the nature of the lesion. A review of these features and how they can be selected is the topic of the next section.

4.3 Features

This section reviews the literature for promising feature candidates and discusses feature selection.

4.3.1 Introduction

Many features could be extracted from the segmented lesion. However, the size of the employed dataset restricts the number of feature candidates that can be reasonably assessed. A minimum sample size of 5 observations per feature is recommended in discriminant analysis [Hair et al., 1998]. This recommendation should to be applied to all features considered during discriminant analysis and not just the selected features. It is therefore crucial to reduce the number of feature candidates beforehand. This section shows two ways to achieve the aim. First, the literature was reviewed for promising features and, second, methods for selecting features in an unsupervised way were discussed.

4.3.2 Feature Candidates

In order to determine these features, the literature of MR breast lesion classification studies was reviewed. Studies were included in this review when the classification was only based on features computed from the images and from the patient’s age; linear discriminant analysis or logistic regression was employed; and each class had more than 10 cases. Table 4.1 provides an overview of the reviewed studies. The individual promising features are described in
Table 4.1: Overview of computerized classification studies for MR breast lesions. Lesions were manually (M) or semi-automatic (SA) segmented. Features were combined with stepwise linear discriminant analysis (LDA) or with stepwise logistic regression analysis (LRA). The classification performance was assessed based on a training set (Train.), an independent test set (Test) or by leave-one-out cross-validation (LOO).

4.3.2.1 Shape Features

The use of morphological features for the interpretation of DCE MR mammography is widely recommended [Harms et al., 1993; Stomper et al., 1995; Nunes et al., 1997b]. The irregular and spiculated pattern is a well known sign of malignancies in X-ray mammography and has also proven to be useful for DCE MR mammography. The aim of most shape features is to quantify the irregularity of the lesions shape.

The following shape features were successfully employed for MR breast lesion classification [Sinha et al., 1997; Gilhuijs et al., 1998; Gilhuijs et al., 2002; Shahar et al., 2002; Sonoda, 2003; Gibbs and Turnbull, 2003; Chen et al., 2004].

<table>
<thead>
<tr>
<th>Reference</th>
<th>Number of benign lesions</th>
<th>Number of malignant lesions</th>
<th>Temporal resolution in seconds</th>
<th>Lesion segmentation</th>
<th>Number of assessed features</th>
<th>Number of combined features</th>
<th>Type of combination</th>
<th>Assessment set</th>
<th>Area under ROC curve</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Mussurakis et al., 1998]</td>
<td>30</td>
<td>64</td>
<td>12</td>
<td>M</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>Train.</td>
<td>-</td>
</tr>
<tr>
<td>[Sinha et al., 1997]</td>
<td>23</td>
<td>20</td>
<td>15</td>
<td>SA</td>
<td>18</td>
<td>10</td>
<td>LDA</td>
<td>Train.</td>
<td>-</td>
</tr>
<tr>
<td>[Gilhuijs et al., 1998]</td>
<td>13</td>
<td>15</td>
<td>90</td>
<td>M</td>
<td>9</td>
<td>2</td>
<td>LDA</td>
<td>LOO</td>
<td>0.96</td>
</tr>
<tr>
<td>[Gilhuijs et al., 2002]</td>
<td>40</td>
<td>40</td>
<td>90</td>
<td>SA</td>
<td>9</td>
<td>4</td>
<td>LDA</td>
<td>LOO</td>
<td>0.95</td>
</tr>
<tr>
<td>[Shahar et al., 2002]</td>
<td>15</td>
<td>18</td>
<td>?</td>
<td>M</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>Train.</td>
<td>-</td>
</tr>
<tr>
<td>[Sonoda, 2003]</td>
<td>20</td>
<td>20</td>
<td>60</td>
<td>M</td>
<td>26</td>
<td>8</td>
<td>LDA</td>
<td>Train.</td>
<td>0.81</td>
</tr>
<tr>
<td>[Gibbs and Turnbull, 2003]</td>
<td>34</td>
<td>45</td>
<td>12</td>
<td>M</td>
<td>22</td>
<td>5</td>
<td>LRA</td>
<td>Test</td>
<td>0.92</td>
</tr>
<tr>
<td>[Chen et al., 2004]</td>
<td>44</td>
<td>77</td>
<td>60</td>
<td>M</td>
<td>14</td>
<td>4</td>
<td>LDA</td>
<td>LOO</td>
<td>0.80</td>
</tr>
</tbody>
</table>

...
4.3 Features

Sonoda, 2003; Chen et al., 2004.

- **Compactness** was defined as $S_{\text{comp}} = S^2/V$, where $S$ and $V$ denote the surface area and the volume of the lesion, respectively. The compactness of a sphere is $6\sqrt{\pi}$. All other shapes have higher compactness values.

- **Spherical Shape Index** was determined by $S_{\text{ssi}} = S/V^{2/3}$. $S_{\text{ssi}}$ relates to $S_{\text{comp}}$ by $S_{\text{ssi}} = S_{\text{comp}}V^{1/3}/\sqrt{S}$. The spherical shape index of a sphere equals $(36\pi)^{1/3}$.

- **Surface Irregularity** was calculated by $S_{\text{sir}} = 1 - S_{\text{sphere}}/S$, where $S_{\text{sphere}}$ denotes the surface area of a sphere of the same volume as the lesion, i.e. a sphere with a radius of $[3V/(4\pi)]^{1/3}$. $S_{\text{sir}}$ is zero for a sphere, and otherwise positive and less than one.

- **Radial Length Ratio** was defined by the ratio of the shortest and the largest radial length, i.e. $S_{\text{rrl}} = \min(S_{rl})/\max(S_{rl})$, where $S_{rl}$ denotes the set of radial lengths for all surface vertices. The radial length of a vertex was determined by its Euclidean distance to the lesion’s centre of gravity. $S_{\text{rrl}} \in (0, 1]$ with spheres having a value of 1.

- **Entropy of Radial Length Distribution** was calculated as $S_{\text{erl}} = -\sum_{n=1}^{100} P_n \log(P_n)$, where $P_n$ stands for the probability that a surface vertex has a radial length that lies within the $n$th increment of the distribution. $S_{\text{erl}} = 0$ for a sphere and $S_{\text{erl}} = 2.0$ for a lesion that has a uniform radial length distribution.

- **Topology Value** was defined by $S_{\text{top}} = 1 - n$, where $n$ is the number of holes within the lesion.

- **Maximum Variance of Radial Gradient Histogram** was calculated by $S_{\text{mvrgh}} = \max_{k=1,...,K}[\var H_{x\in\mathcal{R}}(p_{x,k})]$, where $H_{x\in\mathcal{R}}(p_{x,k})$ stands for the histogram of the normalized radial gradients $p_{x,k}$ for all $x \in \mathcal{R}$,

$$p_{x,k} = \frac{\|\nabla[I(x,k) - I(x,0)] \cdot (x - x_c)\|}{\|\nabla[I(x,k) - I(x,0)]\| \cdot \|x - x_c\|}, \quad (4.3)$$

$x_c$ denotes the centre of region $\mathcal{R}$ and $\nabla[I(x,k) - I(x,0)]$ represents the local image intensity gradient at position $x$ for the difference image of the pre-contrast image from the $k$th post-contrast image. $p_{x,k} = 1$ if the local image intensity gradient at $x$ coincides with the radial direction. Spherical lesions will therefore create a normalized radial gradient histogram that has a peak around one. Arbitrary shapes will result in more uniformly distributed $p_{x,k}$ values and hence in $S_{\text{mvrgh}}$ values closer to zero.
Table 4.2: Classification performance of promising shape features from previous studies. Features that were assessed in this work are marked by √. The performance was measured by the probability that the mean feature value is the same for benign and malignant lesions, and by the area under the ROC curve (AUC). Statistically significant probabilities at the 5% level are marked with a box.

Features included at the \( n \)th step of stepwise discriminant analysis are labelled with \( s^n \).

- **Early Variance of Radial Gradient Histogram** was calculated by
  \[
  S_{\text{evrgh}} = \text{var}_x H_{x \in \mathbb{R}}(p_x, 1),
  \]
  where \( H_{x \in \mathbb{R}}(p_x, 1) \) denotes the radial gradients histogram for the first post-contrast as defined for \( S_{\text{mvrgh}} \).

The performance of the promising shape features are listed in Table 4.2. Of these, \( S_{\text{ssi}} \) and \( S_{\text{sir}} \) were not further assessed, since they were closely related to \( S_{\text{comp}} \). Also excluded were \( S_{\text{ecc}} \) and \( S_{\text{elo}} \), because of their late inclusion during stepwise discriminant analysis. \( S_{\text{evrgh}} \) was not tested due to its similarity with \( S_{\text{mvrgh}} \). Five shape features were further evaluated in Chapter 7, namely \( S_{\text{comp}}, S_{\text{rrl}}, S_{\text{erl}}, S_{\text{top}} \) and \( S_{\text{mvrgh}} \).

### 4.3.2.2 Margin Features

The terms for lesion shape and margin are often merged in the MR literature [Ikeda et al., 2001]. Lobulated might for example stand for a lobulated shape with smooth margins,
while spiculated might describe a lesion of any shape (round or irregular) that has spiculated borders. In comparison, the margin features defined in this section aim to quantify the visibility of the lesion boundary. Margin terms that describe the local shape of lesions, like the spiculated borders, were classified as shape features in this work.

The motivation for assessing the sharpness of the lesion boundary lies in the observation that the infiltration of malignant cells into the surrounding tissue will often lead to ill-defined margins. While in contrast, benign lesions are more likely to have well-defined margins.

The sharpness of the MR breast lesion boundaries was assessed by [Gilhuijs et al., 1998; Gilhuijs et al., 2002; Chen et al., 2004]. The following three margin features were successful in discriminating benign and malignant MR breast lesions.

- **Maximum Margin Gradient** was defined by
  \[
  M_{mmg} = \max_{k=0,1,...,K} \left( \frac{\text{mean}_{x \in B} ||\nabla [I(x,k) - I(x,0)]||}{\text{mean}_{x \in B} I(x,k)} \right), \tag{4.4}
  \]
  where \(\nabla [I(x,k) - I(x,0)]\) denotes the local image intensity gradient at position \(x\) for the difference image \(I(x,k) - I(x,0)\) and \(B\) represents the lesion boundary.

- **Early Margin Gradient** was determined by
  \[
  M_{emg} = \frac{\text{mean}_{x \in B} ||\nabla [I(x,1) - I(x,0)]||}{\text{mean}_{x \in B} I(x,1)}. \tag{4.5}
  \]

- **Variance of Maximum Margin Gradient** was calculated by
  \[
  M_{vmmg} = \frac{\text{var}_{x \in B} ||\nabla [I(x,k_{max}) - I(x,0)]||}{[\text{mean}_{x \in B} I(x,k_{max})]^2}, \tag{4.6}
  \]
  where \(k_{max}\) refers to the image index \(k\) that maximizes equation (4.4).

Table 4.3 lists the performance of these features. Only \(M_{emg}\) was significant when discriminating between benign and malignant lesions, while all three features were included during stepwise discriminant analysis. Note: \(M_{mmg}\) was not tested in this study, because of its closeness to \(M_{emg}\).

### 4.3.2.3 Kinetic Features

The intensity change characteristic due to the injection of a contrast agent is regarded by some as the most important feature for classifying MR breast malignancies [Kaiser and Zeitler, 1989; Stack et al., 1990; Boetes et al., 1994; Gilles et al., 1994; Kelcz et al., 1996]. These so-called kinetic features are motivated by the observation that malignant lesions have generally an increased vascularity and an increased vascular permeability compared to benign...
4.3 Features

Assessed by [Gilhuijs et al., 2002] and [Gilhuijs et al., 1998].

<table>
<thead>
<tr>
<th>Feature</th>
<th>Probability</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{mmg}$</td>
<td>$&gt;0.05$</td>
<td>0.88</td>
</tr>
<tr>
<td>$M_{emg}$</td>
<td>$&lt;0.01$</td>
<td>0.86*2</td>
</tr>
<tr>
<td>$M_{vmmg}$</td>
<td>$&gt;0.05$</td>
<td>0.60*5</td>
</tr>
</tbody>
</table>

Table 4.3: Classification performance of promising margin features from previous studies. Features that were assessed in this work are marked by √. The performance was measured by the probability that the mean feature value is the same for benign and malignant lesions, and by the area under the ROC curve (AUC). Statistically significant probabilities at the 5% level are marked with a box.

Features included at the $n$th step of stepwise discriminant analysis are labelled with $^*n$.

Lesions. Moreover, this increased vascularity and vascular permeability will lead to a faster distribution of the contrast agent into and out of the interstitial space.

The distribution of the contrast agent over time is monitored by acquiring a sequence of images. Ideally, physiological meaningful parameters are derived by fitting a sophisticated pharmacokinetic model to the intensity-time curve, i.e. the mean intensity value of a region of interest measured over time. However, this requires a very high temporal resolution of around 10s, which currently cannot be achieved when imaging the whole breasts at a satisfactory spatial resolution. Assessment relies instead on the general shape of the intensity-time curve, as shown in Figure 1.7 on page 33, and on quantitative measures like the relative initial enhancement and maximum washout [Schnall and Ikeda, 1999; Rankin, 2000; Kuhl and Schild, 2000].

The following notation was used for describing the kinetic features. The time course of the image intensity at voxel location $x$ was denoted by $I(x,k)$ for $k \in \{0,1,2,\ldots,K\}$ for a dynamic sequence with one pre-contrast image and $K$ post-contrast images. The relative time after contrast injection ($t$) and the image index $k$ were related by $t = k\Delta t$, where $\Delta t$ denotes the acquisition time for a single image. The mean intensity of a lesion with region $\mathcal{L}$ at time $t$ was determined by $\bar{I}(t) = \bar{I}_\mathcal{L}(t) = \text{mean}_{x \in \mathcal{L}} I(x,t)$. The standard deviation of the image intensities of lesion $\mathcal{L}$ at time $t$ was denoted by $\sigma_\mathcal{L}(t)$. The maximum lesion intensity was
defined by $\bar{I}_{\text{max}} = \max_t \bar{I}(t) = \bar{I}(t_{\text{max}})$, where $t_{\text{max}}$ denotes the time at which this maximum occurred and $k_{\text{max}}$ stands for the corresponding image index, i.e. $k_{\text{max}} = t_{\text{max}}/\Delta t$.

The kinetic features were generally derived from the mean value for the whole lesion, i.e. $\bar{I}_L(t)$. In contrast, [Chen et al., 2004] calculated the normalized enhancement per voxel, i.e. $I(x, k)/I(x, 0) - 1$, before taking the mean over the region of the lesion. [Mussurakis et al., 1998] obtained features from the central subregion $C$ and the peripheral subregion $P$ of 2D lesion. These subregions were extracted by a line thinning algorithm. $C$ contained the inner 25% and $P$ the outer 25% of the lesion pixels. The actual percentage varied slightly to preserve the shape of the lesion.

[Kelcz et al., 1996; Sinha et al., 1997; Sonoda, 2003] derived kinetic features from fitting $\bar{I}(t)$ to the general saturation equation of the form as follows

$$\bar{I}(t) = I_0 + \left[ \frac{I_a}{(\frac{T_{1/2}}{t})^m + 1} \right],$$

(4.7)

where $I_0$ denotes the mean intensity before contrast administration, $I_a$ represents the asymptotic intensity increase after contrast administration, $T_{1/2}$ stands for the time to half maximum enhancement and $m$ denotes the slope factor. Image sequences with low temporal resolution do not provide enough samples to estimate all four parameters. Instead $I_0$ and $I_a$ can directly be approximated from the data using $I_0 = \bar{I}(0)$ and $I_a = \bar{I}_{\text{max}} - I_0$.

Kinetic features previously employed for computer assisted MR breast lesion classification [Kelcz et al., 1996; Mussurakis et al., 1998; Sinha et al., 1997; Ikeda et al., 1999; Gilhuijs et al., 2002; Sonoda, 2003; Gibbs and Turnbull, 2003; Chen et al., 2004] included

- **Relative Maximum Enhancement**: $K_{rme} = \bar{I}_{\text{max}}/\bar{I}(0) - 1$,
- **Slope Factor**: $K_{slo} = m$,
- **Time to Half Maximum Enhancement (s)**: $K_{bht} = T_{1/2}$,
- **Time to Maximum Enhancement (s)**: $K_{tme} = t_{\text{max}}$,
- **Maximum Rate of Enhancement (1/s)**: $K_{mre} = \max_t \frac{\bar{I}(t + \Delta t)}{\bar{I}(t)} - \bar{I}(0)/\Delta t$,
- **Standard Deviation of Enhancement**: $K_{sde} = \sqrt{\text{var}_t \bar{I}(t)}$,
- **Relative Enhancement at 1 Minute**: $K_{rer1} = \bar{I}(60)/\bar{I}(0) - 1$,
- **Relative Initial Enhancement**: $K_{rie} = \bar{I}(\Delta t)/\bar{I}(0) - 1$,
- **Maximum Washout**: $K_{mwo} = 1 - \bar{I}(K\Delta t)/\bar{I}_{\text{max}}$,
- **General Washout**: $K_{gwo} = [\bar{I}(K\Delta t) - \bar{I}(\Delta t)]/\bar{I}(0)$,
- **Rate of Maximum Washout**: $K_{rmwo} = \frac{\bar{I}_{\text{max}} - \bar{I}(K\Delta t)}{K\Delta t - t_{\text{max}}}$,
4.3 Features

- **Signal Enhancing Ratio:** \( K_{ser} = K_{gwo}/K_{rie} \).

- **Late Peripheral Enhancement:** \( K_{lpe} = \frac{I_p(K \Delta t) - I_c(K \Delta t)}{\sqrt{\sigma_p^2(K \Delta t) + \sigma_c^2(K \Delta t)}} \).

Table 4.4 lists the performance of the promising kinetic features. Benign lesions had on average a significantly lower values for the relative maximum enhancement, the slope factor, the maximum rate of enhancement, the standard deviation of enhancement, the relative enhancement at 1 minute and the maximum washout, while the mean time to maximum enhancement was significantly increased.

Several of these features are likely to be highly correlated. The time to half maximum enhancement \( (K_{ht}) \) was excluded because of its similarity to the time to maximum enhancement \( (K_{tm}) \). The relative enhancement at one minute \( (K_{er1}) \) was employed in previous studies for image sequences with a high temporal resolution. The datasets used in this work had a temporal resolution of 90s. \( K_{er1} \) could therefore only be estimated and would be highly correlated to the relative initial enhancement \( (K_{rie}) \), and hence was omitted. For a temporal resolution of 90s it is also most likely that the maximum rate of enhancement \( (K_{mre}) \) occurs between the pre-contrast and the first post-contrast image. \( K_{mre} \) is then equal to \( [K_{rie} - 1]/\Delta t \) and was therefore not assessed in this study.

### 4.3.2.4 Texture Features

The appearance of the internal enhancement is another feature of radiological interest for MR breast lesions [Ikeda et al., 2000]. Radiologists describe these with terms like homogeneous, heterogeneous, stippled and clumped, see Table 1.1, page 34. Additional important features are rim enhancement and septations, which are present on a coarser image scale [Abdolmaleki et al., 1997; Nunes et al., 1997a]. Homogeneous enhancements and nonenhancing septations are associated with benign disease. Rim enhancement is often a sign of malignancy, since a lesion with a necrotic centre will always have this appearance.

This section reviews features that were employed to capture the fine scale appearance, i.e., texture, of MR breast lesions. Late peripheral enhancement \( (K_{lpe}) \), a coarse scale feature employed for detecting rim enhancements, was described in Section 4.3.2.3. The texture of MR breast lesions was quantified by two approaches. The first employed statistical quantities derived from spatial grey level dependence matrices [Sinha et al., 1997; Sonoda, 2003; Gibbs and Turnbull, 2003]. The second was based on the variance of the image intensities within the lesion [Gilhuijs et al., 1998; Chen et al., 2004].

The use of spatial grey level dependence matrices to describe image textures was proposed
## 4.3 Features

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta t$ (s)</td>
<td>16</td>
<td>12</td>
<td>15</td>
<td>60</td>
<td>90</td>
<td>60</td>
<td>12</td>
<td>60</td>
</tr>
<tr>
<td>Feature</td>
<td>Probability</td>
<td>AUC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{rme}$</td>
<td>$\sqrt{0.018}$</td>
<td>0.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{slf}$</td>
<td>$&lt;0.001$</td>
<td>$0.000$</td>
<td>$0.033$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{hit}$</td>
<td>$0.094^8$</td>
<td>0.370</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{tmx}$</td>
<td>$&lt;0.01^{84}$</td>
<td>$0.000^{81}$</td>
<td>0.75$^{84}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{mre}$</td>
<td>$0.000^{82}$</td>
<td>0.159</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{sde}$</td>
<td>$0.008$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{rer1}$</td>
<td>$0.001^p$</td>
<td>0.018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{rie}$</td>
<td>$&lt;0.05$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{mwo}$</td>
<td>$&lt;0.01^{82}$</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{gwo}$</td>
<td>$&lt;0.001^{41}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{rmwo}$</td>
<td>0.73$^{83}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{ser}$</td>
<td>$&lt;0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K_{spe}$</td>
<td>$0.001$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.4: Classification performance of promising kinetic features from previous studies. Features that were assessed in this work are marked by $\sqrt{\text{✓}}$. The temporal resolution of the contrast-enhanced image sequence is stated in the row labelled with $\Delta t$. The classification performance was measured by the probability that the mean feature value is the same for benign and malignant lesions, and by the area under the ROC curve (AUC). Statistically significant probabilities at the 5% level are marked with a $\text{box}$. Features included at the $n$th step of stepwise discriminant analysis are labelled with $^{sn}$.

by [Haralick et al., 1973]. The $(i,j)$th entry of the spatial grey level dependence matrix $P_{d}^{\alpha\beta}$ records the probability that there exists a voxel with intensity $j$ at a distance $d$ in direction $\alpha\beta$ from a voxel with intensity $i$. The direction $\alpha\beta$ is defined by the angle $\alpha$ in the $x_1x_2$-plane and the angle $\beta$ in the $x_1x_3$-plane.

The following quantities, derived from the spatial grey level dependence matrix, were suc-
features successfully employed in previous studies [Sinha et al., 1997; Sonoda, 2003; Gibbs and Turnbull, 2003].

- **Angular Second Moment:** \(T_{asm} = \sum_i \sum_j P(i, j)^2\),
- **Contrast:** \(T_{con} = \sum_{k=0}^{K-1} k^2 P_{a-b}(k)\),
- **Variance:** \(T_{var} = \sum_i (i - \mu_a)^2 P_a(i)\),
- **Inverse Difference Moment:** \(T_{idm} = \sum_i \sum_j \frac{1}{1+(i-j)^2} P(i, j)\),
- **Sum Average:** \(T_{sua} = \sum_{k=2}^{2K} k P_{a+b}(k)\),
- **Sum Entropy:** \(T_{sue} = - \sum_{k=1}^{2K} P_{a+b}(k) \log(P_{a+b}(k))\),
- **Entropy:** \(T_{ent} = - \sum_i \sum_j P(i, j) \log(P(i, j))\),
- **Difference Entropy:** \(T_{die} = - \sum_{k=0}^{K-1} P_{a-b}(k) \log(P_{a-b}(k))\),

where \(K\) denotes the number of grey levels, \(P_a(i) = \sum_j P(i, j)\), \(\mu_a = \sum_i i P_a(i)\), \(P_{a+b}(k) = \sum_i \sum_{j=|k-i|} P(i, j)\) and \(P_{a-b}(k) = \sum_i \sum_{|i-j|=k} P(i, j)\).

Texture features were either extracted from the 2D regions of the post-contrast images [Sinha et al., 1997], or from the core of 3D segmentations of the difference images [Sonoda, 2003], or from 2D regions of the fat-suppressed post-contrast images [Gibbs and Turnbull, 2003]. The image intensities were either reduced to 28 [Sinha et al., 1997], or 40 [Sonoda, 2003] or 32 [Gibbs and Turnbull, 2003] grey levels. [Gibbs and Turnbull, 2003] employed a histogram equalization before grey level reduction. All studies created spatial grey level dependence matrices for a distance \(d\) of one pixel (voxel) and for the directions \(\alpha\) (and \(\beta\)) \(\in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}\). Texture features were then derived from the average of these matrices.

The group from the University of Chicago employed enhancement-variance features to quantify the inhomogeneity of contrast uptake within the lesion [Gilhuijs et al., 1998; Chen et al., 2004]. They defined the maximum uptake variation by \(T_{muv} = \max_{k=0,\ldots,K} V(k)/V(0)\), where \(V(k) = \text{var}_{x \in L} I(x, k)\). \(T_{muv} \geq 1\) and records the maximum increase in intensity variation within the lesion relative to the pre-contrast image. Of the features that were based on \(T_{muv}\), the following two features were included during stepwise linear discriminant analysis.

- **Index of Maximum Uptake Variation** was calculated by \(T_{imuv} = k_{muv}\) such that \(V(k_{muv})/V(0) = T_{muv}\).
- **Increase Rate of Maximum Uptake Variation** was defined by \(T_{irmuv} = T_{muv}/T_{imuv}\).

The performance of the promising texture features is listed in Table 4.5. It can be observed that the features with the lowest class discrimination probability were not necessarily included during stepwise discriminant analysis. Malignant lesions had on average higher entropy and
### Table 4.5:

Classification performance of promising texture features from previous studies. Features that were assessed in this work are marked by √. The performance was measured by the probability that the mean feature value is the same for benign and malignant lesions, and by the area under the ROC curve (AUC). Statistically significant probabilities at the 5% level are marked with a box.

Features included at the nth step of stepwise discriminant analysis are labelled with sn.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Probability</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasm</td>
<td>0.019 0.292</td>
<td>0.876 0.049</td>
</tr>
<tr>
<td>Tcon</td>
<td>0.045 0.064</td>
<td>0.049 0.74</td>
</tr>
<tr>
<td>Vars</td>
<td>0.055 0.061</td>
<td>0.503 0.58</td>
</tr>
<tr>
<td>Tidm</td>
<td>0.009 0.007</td>
<td>0.283 0.130</td>
</tr>
<tr>
<td>Tuis</td>
<td>0.073 0.007</td>
<td>0.024 0.045</td>
</tr>
<tr>
<td>Tenti</td>
<td>0.100 0.100</td>
<td>0.436 0.74</td>
</tr>
<tr>
<td>Timu</td>
<td>0.009 0.007</td>
<td>0.283 0.130</td>
</tr>
<tr>
<td>Tirm</td>
<td>0.073 0.007</td>
<td>0.024 0.045</td>
</tr>
</tbody>
</table>

### variances values as well as lower angular second moments than benign lesions [Sinha et al., 1997; Sonoda, 2003; Gibbs and Turnbull, 2003]. This indicates that malignant lesions had a more heterogeneous appearance.

Only texture features that achieved a class discrimination probability below 0.05 or features that were included during the first 3 steps of discriminant analysis will be assessed further in this work.

### 4.3.2.5 Other Features

This section describes features that do not fall into any of the previous feature categories.

[Gibbs and Turnbull, 2003] reported that the mean lesion size was statistically significantly greater, at the 0.5% level, for malignant lesions. Their logistic regression model included
lesion size. Lesion size ($O_{volume}$) was therefore further assessed.

An image classification system should, by definition, only rely on features extracted from the images. However, the final classification rule of [Sinha et al., 1997; Gibbs and Turnbull, 2003] also incorporated the patient’s age. Patients with malignant lesions were on average significantly older, at the 0.7% level, in [Sinha et al., 1997]. To allow for a comparison, patient age ($O_{age}$) was therefore also included in this work.

4.3.2.6 Summary

This section reviewed quantitative image features, which were previously successfully employed for the discrimination of MR breast lesions. The aim was to determine a set of satisfactory image features that can automatically be derived from segmented lesions. The exception was patient age, which was included to allow comparison with previous studies.

The size of the dataset constrains the number of features that can be evaluated. Therefore less effective features and obviously highly correlated features were excluded. The final set consisted of 25 features as summarized in Table 4.6. A set of 25 features requires a minimum of 125 observations for linear discriminant analysis [Hair et al., 1998]. Strategies to reduce the number of feature candidates for smaller datasets are discussed in the next section.

4.3.3 Unsupervised Feature Selection

The limitation on the number of feature candidates for linear discriminant analysis includes not only the selected features, but all features that were assessed [Hair et al., 1998]. Generally, all features that are evaluated with respect to their classification performance, i.e. in a supervised way, will count as feature candidates. The difficult task is therefore to find irrelevant or redundant features without knowing the class assignment.

Without knowledge of the class assignment, relevance can be estimated only indirectly. Assuming that cases from the same class have more similar feature values than cases from different classes, features could be selected on the basis of their ability to uncover clusters [Dy and Brodley, 2004]. Various similarity measures and cluster criteria exist [Duda et al., 2001], and the problem remains that of choosing a suitable one for the given application.

Redundancy is normally measured by the correlation between features. A feature that is highly correlated to an already selected feature, adds little to the discrimination power and will therefore not be selected during multiple discriminant analysis. Removing highly correlated features before multiple discriminant analysis seems therefore a good option. Drawbacks stem from the assumed linearity and the possible loss of the optimum features with regard to class discrimination.
### 4.3 Features

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Category</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{\text{comp}}$</td>
<td>Shape</td>
<td>Compactness</td>
<td>132</td>
</tr>
<tr>
<td>$S_{\text{rrl}}$</td>
<td>Shape</td>
<td>Radial length ratio</td>
<td>132</td>
</tr>
<tr>
<td>$S_{\text{erl}}$</td>
<td>Shape</td>
<td>Entropy of radial length distribution</td>
<td>132</td>
</tr>
<tr>
<td>$S_{\text{top}}$</td>
<td>Shape</td>
<td>Topology value</td>
<td>132</td>
</tr>
<tr>
<td>$S_{\text{mvrg}}$</td>
<td>Shape</td>
<td>Maximum variance of radial gradient histogram</td>
<td>132</td>
</tr>
<tr>
<td>$M_{\text{emg}}$</td>
<td>Margin</td>
<td>Early margin gradient</td>
<td>134</td>
</tr>
<tr>
<td>$M_{\text{vmmg}}$</td>
<td>Margin</td>
<td>Variance of maximum margin gradient</td>
<td>134</td>
</tr>
<tr>
<td>$K_{\text{rme}}$</td>
<td>Kinetic</td>
<td>Relative maximum enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{stk}}$</td>
<td>Kinetic</td>
<td>Slope factor</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{tmx}}$</td>
<td>Kinetic</td>
<td>Time to maximum enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{sde}}$</td>
<td>Kinetic</td>
<td>Standard deviation of enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{rie}}$</td>
<td>Kinetic</td>
<td>Relative initial enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{mwo}}$</td>
<td>Kinetic</td>
<td>Maximum washout</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{gw}}$</td>
<td>Kinetic</td>
<td>General washout</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{rmwo}}$</td>
<td>Kinetic</td>
<td>Rate of maximum washout</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{ser}}$</td>
<td>Kinetic</td>
<td>Signal enhancing ratio</td>
<td>138</td>
</tr>
<tr>
<td>$K_{\text{lpe}}$</td>
<td>Kinetic</td>
<td>Late peripheral enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$T_{\text{asm}}$</td>
<td>Texture</td>
<td>Angular second moment</td>
<td>139</td>
</tr>
<tr>
<td>$T_{\text{var}}$</td>
<td>Texture</td>
<td>Variance</td>
<td>139</td>
</tr>
<tr>
<td>$T_{\text{sue}}$</td>
<td>Texture</td>
<td>Sum average</td>
<td>139</td>
</tr>
<tr>
<td>$T_{\text{ent}}$</td>
<td>Texture</td>
<td>Entropy</td>
<td>139</td>
</tr>
<tr>
<td>$T_{\text{imuv}}$</td>
<td>Texture</td>
<td>Index of maximum uptake variance</td>
<td>139</td>
</tr>
<tr>
<td>$T_{\text{irmuv}}$</td>
<td>Texture</td>
<td>Increase rate of maximum uptake variance</td>
<td>139</td>
</tr>
<tr>
<td>$O_{\text{volume}}$</td>
<td>Other</td>
<td>Lesion size</td>
<td>140</td>
</tr>
<tr>
<td>$O_{\text{age}}$</td>
<td>Other</td>
<td>Patient age</td>
<td>141</td>
</tr>
</tbody>
</table>

Table 4.6: Overview of features that were previously successfully employed for MR breast lesion classification and that were assessed in this work.

Both approaches are inferior to the optimal classifier because of the unknown class labelling. The use of prior knowledge of the problem domain to obtain more informative features is instead recommended [Duda et al., 2001]. This advice was followed in this thesis by identifying promising features from previous studies. If the number of features need
to be reduced even further, then the removal of highly correlated features is an attractive unsupervised method.

4.3.4 Summary
Twenty-five promising features for MR breast lesion classification were identified in this section, see Table 4.6. Removal of highly correlated features is another option for further reducing the number of feature candidates.

Having selected a suitable number of feature candidates, the next task is to combine these to create an optimal classifier. This will be the topic of the next section.

4.4 Classification
This section reviews classification methods based on multiple discriminant analysis as well as measures and procedures for assessing the classification performance.

4.4.1 Introduction
The general problem of building a classifier based on a training dataset can be formulated as follows. Given a training dataset $\mathcal{T}$ of $N$ cases with $P$ features per case and known class affiliation $c_n$, i.e. $\mathcal{T} = \{x_{pn}, c_n\}$ for $p \in \{1, 2, \ldots, P\}$ and $n \in \{1, 2, \ldots, N\}$, what is the class of a new case? To answer this question, it is necessary to generalize from the given data to the whole $P$-dimensional feature space.

Various approaches exist to create such class models. The class model can often be described by a so-called discrimination function that defines the decision boundary between the classes. The form of this function might be motivated by either the expected data distribution, or by the assumed interpolation function, or by a certain optimization measure, or by the need for fast computation or by the aim to extract rules. In all cases, the discrimination function should not be too flexible, otherwise all training examples are remembered and generalization for unseen data will be poor.

A full review of classification methods is beyond the scope of this thesis. More information, however, can be found in [Duda et al., 2001; Haykin, 1998; Theodoridis and Koutroumbas, 1999]. Linear discriminant analysis and logistic regression are two of the most widely used classification methods. These two methods will be reviewed in the next two sections. These reviews will be based on [Hair et al., 1998; Hosmer and Lemeshow, 2000].

4.4.2 Linear Discriminant Analysis
Discriminant analysis is a statistical method with the basic purpose to estimate the relationship between a set of metric independent variables, i.e. the features, and a single categorical
dependent variable. The dependent variable may take values from two or more groups or classes. The aim of classification is to predict the group of a sample from the values of its independent variables.

In linear discriminant analysis, the features are linearly combined using the following discriminant function

$$z_n = w_0 + \sum_{p=1}^{P} w_p x_{pn},$$

(4.8)

where $z_n$ stands for the discriminant score for sample $n$, $w_0$ denotes the free parameter for the intercept, $w_p$ represents the free weight parameter for feature $p$, and $x_{pn}$ denotes the value of feature $p$ for sample $n$.

Optimal discrimination of the data from two classes, namely $A$ and $B$, is based on maximizing the objective function as follows

$$\max_w D^2 = \frac{(\bar{z}_A - \bar{z}_B)^2 (N_A + N_B - 2)}{(N_A - 1)\sigma_A^2 + (N_B - 1)\sigma_B^2},$$

(4.9)

where $N_A$ denotes the number of samples in class $A$, $\bar{z}_A$ represents the mean of $z_n$ for $n \in A$, and $\sigma_A^2$ stands for the variance of $z_n$ for $n \in A$. $D^2$ is the ratio of the between-class estimate of variance and the within-class estimate of variance, and is known as the Mahalanobis distance. If $\bar{z}_A < \bar{z}_B$, then the optimal classification rule for observation $x_n$ is given by

$$x_n \in A \quad \text{if} \quad z_n < \theta \quad \text{where} \quad \theta = \frac{N_A \bar{z}_B + N_B \bar{z}_A}{N_A + N_B}.$$

(4.10)

The classification performance for the training set can be improved by increasing the size of the feature set. However, the classification performance for an independent test set will only initially improve and then worsen as more features are added. Linear discriminant analysis employs therefore the strategy to select the smallest feature subset that does not substantially reduce the classification performance for the training set.

The contribution of a subset of $m$ features to the prediction accuracy of a full model that includes the $m$ features can be assessed statistically with the partial F-test [Hair et al., 1998]. The F-test assesses the difference of two population variances. In the case of discriminant analysis, the unexplained variances after fitting two models are compared. For the null-hypothesis $H_0$: “the reduced and the full model have the same error variance” and for the alternative hypothesis $H_1$: “the full model has a reduced error variance”, the test statistic value is given by

$$f = \frac{SSE_{P-m} - SSE_P}{\frac{SSE_P}{N - P - 1} \frac{N - P - 1}{m}}.$$

(4.11)
4.4 Classiﬁcation

where $SSE_{P-m}$ and $SSE_P$ denote the error sum of squares, i.e. $SSE = \sum_n^N (c_n - z_n)^2$, for the reduced and full model, respectively [Devore, 1995; Hair et al., 1998]. If $H_0$ is true, then $f$ follows a F distribution with $m$ and $N - P - 1$ degrees of freedom. $H_0$ is rejected, i.e. the full model is needed, for values of $f$ that are too large to have occurred by chance.

Theoretically, all possible subsets of features could be compared in this manner. Practically, however, the number of possible feature subsets may prevent such an exhaustive comparison. Instead, stepwise forward selection (backward elimination) is employed, where one feature is included (excluded) at each step after starting with a model based on no (all) features. Forward selection and backward elimination are often combined, by ﬁrst adding one feature and then assessing if any of the included features became insigniﬁcant.

The key assumptions of discriminant analysis are multivariate normality of the feature values and equal variances and covariances of the feature values across the classes. Violation of these assumptions may cause problems and examination of the data is recommended. Logistic regression, discussed in the next section, can be used as an alternative if these assumptions are violated.

4.4.3 Logistic Regression

Logistic regression is an attractive alternative to discriminant analysis whenever the dependent variable has only two classes. It is more robust than discriminant analysis to violations of the multivariate normality assumption and the equal variance/covariance assumption [Hosmer and Lemeshow, 2000].

Logistic regression is based on a S-shape function $\pi$, which is used to predict the class probability directly, i.e. $\pi(x_n) = P(c_n = 1|x_n)$ for $c_n \in \{0, 1\}$. Note that $P(c_n = 0|x_n)$ is equivalent to $1 - \pi(x_n)$ since it is a two class problem. Function $\pi$ is bounded between 0 and 1, and is deﬁned by

$$
\pi(x_n) = \frac{e^{g(x_n)}}{1 + e^{g(x_n)}},
$$

(4.12)

where $g(x_n) = w_0 + w_1 x_{1n} + w_2 x_{2n} + \ldots + w_P x_{Pn}$. Assuming $N$ independent observations, the log-likelihood function $L(w)$ of the data is given by

$$
L(w) = \ln [P(x_n|c_n)] = \ln \left[ \prod P(c_n = 1|x_n)^{c_n} P(c_n = 0|x_n)^{1-c_n} \right]
$$

(4.13)

$$
= \sum_{n=1}^N \left[ c_n \ln(\pi(x_n)) + (1 - c_n) \ln(1 - \pi(x_n)) \right].
$$

The objective in logistic regression is to choose a parameter vector $w$ such that $L(w)$ is maximized. This so-called maximum likelihood estimator, $\hat{w}$, can be estimated by differentiating $L(w)$ with respect to $w$. 

Similar to linear discriminant analysis, the improvement of the model by adding $m$ features can be statistically assessed. In this case, the test statistic is based on the log likelihood values of the different models, i.e. $g = -2(L_{P-m} - L_P)$ where $L_P$ is the log likelihood of a model with $P$ features. $g$ follows a chi-square distribution with $m$ degrees of freedom if the reduced and the full model have the same log likelihood values. Parsimonious models can then be built using the stepwise inclusion/exclusion technique as for linear discriminant analysis.

While logistic regression provides clear advantages over linear discriminant analysis, the question remains whether logistic regression models are flexible enough. Multi-layer neural networks with logistic activation functions can be considered as a non-linear generalization of logistic regression. However, a review based on 44 papers found no statistical significant difference between both methods in 68% of all papers [Dreiseitl and Ohno-Machado, 2002]. Increased flexibility is also likely to lead to overfitting in small datasets. Interpretability of the model parameter and easier model-building are further advantages of logistic regression in comparison to multi-layer neural networks. Logistic regression was therefore the preferred classification approach employed in this work.

### 4.4.4 Evaluation Measures and Techniques

This section reviews measures and techniques used for evaluating the classification performance.

#### 4.4.4.1 Classification Table

The classification results for a fixed decision boundary can be summarized using a classification table as shown in Table 4.7. For two classes, the classification results can fall into one of four categories, namely true positive, false positive, false negative and true negative. The table entries state the number of cases that fall into each category.

<table>
<thead>
<tr>
<th>Predicted</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malignant</td>
<td>Malignant</td>
</tr>
<tr>
<td>Benign</td>
<td>Malignant</td>
</tr>
</tbody>
</table>

Table 4.7: Classification table for the two class problem of benign and malignant cases.

The following measures can be derived from the classification table:

- **Sensitivity aka True positive rate:** $\frac{TP}{TP+FN}$,
- **Specificity aka True negative rate:** $\frac{TN}{TN+FP}$,
- **Accuracy:** $\frac{(TP+TN)}{(TP+FP+FN+TN)}$,
with all measures taking values within $[0,1]$. The comparison of these measures for different studies is limited because different penalties might have been used for the misclassification of malignant lesions versus the misclassification of benign lesions, which leads to different trade-offs between sensitivity and specificity. Receiver operator characteristics analysis, which is discussed next, enables the comparison of different studies regardless of their favoured optimization criteria.

### 4.4.4.2 Receiver Operating Characteristics Analysis

By tradition, the receiver operating characteristics (ROC) curve is a plot with the false positive rate, i.e. 1-specificity, on the x-axis and the true positive rate, i.e. sensitivity, on the y-axis. The curve is created by varying the thresholds on the decision criterion and provides a visual tool to examine the trade-off between sensitivity and specificity. The point $(0,1)$ represents a perfect classification, while the point $(1,0)$ stands for an incorrect classification of all cases. The ROC curve will always go through the point $(0,0)$ where all cases are classified as benign and through the point $(1,1)$ where all cases are classified as malignant.

The classification performance can be summarized by the area under the ROC curve ($AUC$), which represents the probability that a randomly chosen malignant lesion is rated more suspicious than a randomly chosen benign lesion [Hanley and McNeil, 1982]. Assuming binormal distributions, a smooth curve can be fitted to the points of a ROC curve using a maximum likelihood estimation [Metz et al., 1998]. The statistical significance of the difference between ROC curves can then be tested based on the areas under the fitted ROC curves [Metz et al., 1998].

### 4.4.4.3 Leave-One-Out Cross-Validation

Assessing the classification performance based on the training data will only show how well the discrimination function can fit the data, but will not provide information regarding its generalization performance. Ideally, the classification performance should be assessed on a test set that is independent of the training set. Furthermore, both datasets should comprise enough samples to provide reliable results. In reality, data collection is often very expensive and datasets are therefore most of the time relatively small.

A better technique in this case is leave-one-out cross-validation, which keeps the independence between the training and test set, while making maximum use of the available data. For $N$ samples, the classifier is trained on $N-1$ samples and tested on the left-out sample. This is repeated $N$ times, where each time a different sample is left out. The classification error probability is finally estimated from the total number of errors for the $N$ tests. Leave-
one-out ROC curves can be generated by using the ratio of the a posteriori probabilities, i.e. $P(c_n = 1|x_n)/P(c_n = 0|x_n)$, of the left-out case as the classifier result for the Nth test, and combining these results to a single ROC curve. For statistically independent samples, the leave-one-out method provides an unbiased estimate of the average classification error probability, for a classifier designed using a finite set of N training samples, over all possible training sets of size N [Theodoridis and Koutroumbas, 1999]. The major drawback of the leave-one-out method is its computational complexity.

### 4.4.5 Summary

This section reviewed the most common classification methods for MR breast lesions as well as measures and techniques to evaluate the classification performance. In conclusion, logistic regression seems a good choice as a classification method for relatively small datasets, and the leave-one-out ROC curve analysis is recommended for assessing the classification performance.

### 4.5 Summary

This chapter reviewed automatic segmentation and classification methods for MR breast lesions. The aim was to identify and define a suitable framework for the computer aided diagnosis system for DCE MR mammography described in Chapter 7.

Only a few segmentation methods have been proposed for MR breast lesions and these often relied on clustering the intensity-time curves without considering the spatial relationship of the voxels within a cluster. The author therefore decided to develop a semi-automatic segmentation method, which extracts the most plausible lesion with respect to the statistics provided by a crude outline.

The number of feature candidates when creating a classifier needs to be restricted to avoid overfitting and bias of the leave-one-out performance estimate. The literature was therefore systematically reviewed and 25 promising features, for discriminating between benign and malignant MR breast lesions, were identified and defined. Removal of the most correlated features was recognized as a suitable method for further reducing the number of features. Two common classifier models, namely linear discriminant analysis and logistic regression, were discussed and logistic regression was selected as the favourite model because of its greater robustness to violations of statistical assumptions. Finally, measures and techniques for evaluating the classification performance were reviewed and ROC curve analysis for leave-one-out cross-validation was recommended.

In conclusion, a semi-automatic segmentation approach which extracts the most plausible
lesion, the least correlated features from 25 feature candidates (Table 4.6), logistic regression and evaluation based on leave-one-out ROC curves were selected as a suitable framework for the work which will be described in Chapter 7.
Chapter 5

Comparison of Biomechanical Breast Models

This chapter describes an evaluation study of the accuracy with which biomechanical breast models can predict the displacements of tissue within the breast. The aim of the study was to determine a suitable model configuration, which can be employed for simulating plausible breast deformations for the validation of a registration algorithm in Chapter 6. The content of this chapter is based on and extends [Tanner et al., 2001; Tanner et al., 2002a]

5.1 Introduction

This chapter describes a systematic evaluation of biomechanical breast models for two data sets. The evaluation was based on the acquisition of two breast MR image volumes of two volunteers, with compression plates in two different positions providing two different deformations. The surface displacements of the breasts were calculated by applying a full 3D non-rigid registration to the high resolution 3D pre- and post-deformed images. These surface displacements were then used as a boundary condition for the FEM model. The result of solving the FEM model was evaluated on the basis of the residual displacement error of manually identified point landmarks. Experiments were conducted for six linear and three non-linear material models for a lateral compression of about 20%. These experiments were repeated for a range of different surface boundary conditions, Poisson’s ratios and skin properties. The influence of the mesh resolution and the optimization strategy was assessed with one dataset. The effectiveness of transverse isotropic material models and hyperelastic material models was tested. The performance of finite and infinitesimal deformation formulations were compared.

This is the first study that assessed the influence of all these aspects on the accuracy of the biomechanical breast models. Especially novel is the comparison of Mooney-Rivlin hyperelastic models (employing a finite deformation formulation) with simpler approaches, and the use of transverse isotropic material models for predicting breast deformations.

The rest of the chapter is structured as follows. Section 5.2 describes the biomechanical models, the evaluation method and the test objectives and test configurations. The results of the study are presented in Section 5.3. The main findings are summarized and discussed
5.2 Materials and Methods

This section provided information about the design of the biomechanical breast models and the tested configurations.

5.2.1 Datasets

Two volunteers were recruited for this study, a 37 year old nulliparous pre-menopausal women and a 65 year old multiparous post-menopausal women on HRT. T1 weighted images of the volunteers were acquired on a Philips 1.5T Intera using a 3D fast gradient echo sequence with TR = 16.9, TE = 6.0, flip angle 35°, axial slice direction, and a spatial resolution of 0.82×0.82×2mm and 0.82×0.82×2.5mm, respectively. Each volunteer was positioned in a fixation device for breast biopsies provided by Philips Medical Systems. The right breast was placed between two plates, in the sagittal plane, without any compression for the first image. For the second image, the plate on the breast’s medial side was kept immobile, while the plate on the lateral side was moved manually towards the immobile plate by as much as the volunteer could comfortably tolerate. The images showed that this resulted in a compression of about 20% in breast thickness for each volunteer (from 75mm to 60mm for volunteer one and from 78mm to 63mm for volunteer two). Volunteers were asked not to move between these two scans. Consent of volunteers were obtained in accordance with Guy’s Hospital local research ethical approval 00/11/99. Example slices of these images are shown in Figure 5.12(a) and 5.13(a) on page 167 and 168.

5.2.2 Non-rigid Registration

Surface displacements were derived from a full 3D non-rigid registration, rather than a 3D surface registration [Ferrant et al., 1999], to improve accuracy by avoiding segmentation and decimation errors. The non-rigid registration algorithm described in [Rueckert et al., 1999b], and reviewed in Chapter 2 (see page 80), was used to register the image of the compressed breast to the image of the uncompressed breast. The registration was not constrained to conserve volume because blood volume may have been reduced due to the compression. The registration algorithm was applied in a multi-resolution fashion. After an initial rigid registration, the images were non-rigidly registered using a regular grid of control points approximated by B-splines. The initial control point spacing was 40mm. At every resolution level this distance was halved. Based on a visual assessment, i.e. when the images looked sufficiently aligned, the registration was stopped at a final control point spacing of 10mm for
volunteer one and 2.5mm for volunteer two. The fact that both images were acquired within 15 minutes, in the same setting, with a high image resolution and without the administration of contrast agent simplified the task of registration. Example slices of the registered images are shown in Figure 5.12(b) and 5.13(b).

5.2.3 Biomechanical Breast Models

This section provides details on the design and the configuration of the tested biomechanical breast models.

5.2.3.1 Geometric FEM Model

This part describes the creation of the finite element meshes from the MR images.

The 3D MR images were firstly corrected for intensity nonuniformity using N3 from [Sled et al., 1998], which assumes a smooth, slowly varying, multiplicative bias field with Gaussian intensity distortions. The images of the uncompressed breast were then manually segmented into breast and background, using an interactive tool in ANALYZE [AnalyzeDirect Inc., 2005]. The pectoral muscle was included in the background. The breast regions were further segmented into fatty and fibroglandular tissue using thresholding followed by manual editing when necessary. The breast segmentations were smoothed by blurring them using an isotropic 3D Gaussian kernel with a standard deviation of 1mm. The blurred segmentations were then down-sampled to an isotropic voxel size of 7.5mm to 10mm (Table 5.1) such that different mesh resolutions could be tested and the limit of available elements was not exceeded during the meshing stage. A 3D triangulation of the outer surface of the breast was obtained using marching cubes, smoothing and decimation techniques from The Visualization Toolkit [Schroeder et al., 1998]. The triangulated volumes were meshed into 10-noded

<table>
<thead>
<tr>
<th>Preprocessing</th>
<th>Volunteer One</th>
<th>Volunteer Two</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isotropic voxel size (mm)</td>
<td>8 8</td>
<td>10 7.5</td>
</tr>
<tr>
<td>Number of decimation steps</td>
<td>120 2</td>
<td>1 1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Mesh configuration</th>
<th>Volunteer One</th>
<th>Volunteer Two</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of nodes</td>
<td>51072 102102</td>
<td>58614 117347</td>
</tr>
<tr>
<td>Number of elements</td>
<td>34873 72756</td>
<td>40636 81633</td>
</tr>
<tr>
<td>Badly shaped elements (%)</td>
<td>0.057 0.003</td>
<td>0.000 0.004</td>
</tr>
</tbody>
</table>

Table 5.1: Preprocessing parameters and geometric FEM model configurations. All triangulations were smoothed by 120 iterations. Elements that exceeded an aspect ratio of 20 or an angle of 165° were classified as 'badly shaped' by ANSYS.
5.2 Materials and Methods

Figure 5.1: Geometric models for volunteer one. (a-d) 2D axial example slice of the 3D MR image or biomechanical breast models at the same position showing (a) MR image of the uncompressed breast, (b) segmentation of (a), (c-d) material assignment to the elements of (c) GM1 and (d) GM2. (e-f) Surface rendering of the undeformed 3D mesh of (e) GM1 and (f) GM2.

tetrahedral elements employing the Delaunay technique mesher from the ANSYS FEM package [ANSYS Inc., 2000]. These elements have an additional node on each edge, which allows for quadratic shape interpolation as described in Section 3.4.1. The skin was modelled by adding 1mm thick triangular prisms with additional nodes in the middle of each edge on the surface of the breast.

Two geometrical models, called GM1 and GM2, were created for each volunteer to investigate the combined influence of the element size and shape. The GM1 models had about half as many elements as the GM2 models, see Table 5.1. This was achieved either by decimating the triangular surface more (volunteer one) or by down-sampling the blurred segmentations more (volunteer two). Figure 5.1 (5.2) shows a 2D example slice of the breast image, the segmentation and the corresponding geometric FEM models for volunteer one (two).

The accuracy of the mesh is a function of the largest dimension of the individual elements, the polynomial order of the shape function and the type of element [Bathe, 1996]. Quadratic tetrahedral elements were reported to be as accurate as linear hexahedral elements when the tetrahedral mesh had the same or twice the number of elements than hexahedral mesh [Cifuentes and Kalbag, 1992; Viceconti et al., 1998]. Voxel based meshes require a large number of elements to achieve a moderate accuracy [Viceconti et al., 1998].

In this work, tetrahedral elements were chosen because of their flexibility to generate unstructured meshes. To improve accuracy, tetrahedrons with a quadratic shape function
5.2 Materials and Methods

Figure 5.2: Geometric models for volunteer two. Axial example slice of (a) the MR image, (b) the segmentation and (c-d) the material assignment to the mesh elements of (c) GM1 and (d) GM2. (e-f) Surface rendering of the mesh of (e) GM1 and (f) GM2.

were employed, a relative fine mesh was used and badly shaped elements were kept to a minimum. In comparison to previously reported studies, the created tetrahedral meshes had at least 2.3 times more elements than the voxel based mesh and were at least 7 to 10 times denser than the surface based hexahedral meshes, see Table 3.3 on page 119.

5.2.3.2 Material Models

As discussed in Section 3.5, published values of the stress to strain relationship of breast tissue types vary considerably. Models were constructed that covered the range of values and complexity of published stress-strain relationships in order to determine their influence on predicted deformations in a systematic and quantitative way.

Linear and Piecewise-Linear Models: The individual tissue types were modelled as isotropic and homogeneous. Only displacement boundary conditions were applied, and hence only the ratios of the Young’s modulus of the different tissue types were required. Six linear models were constructed where fibroglandular tissue was 1, 1.5, 5, 10, 15 or 20 times stiffer than fatty tissue. Three non-linear models were created in accordance to [Krouskop et al., 1998; Azar et al., 2000; Samani et al., 2001b]. These nine models were tested without skin (MM1-MM9) and with skin that was 10 times stiffer than the fatty tissue in the linear models (MM1S-MM9S), see Test4. The non-linear stress-strain functions were approximated by piecewise-linear functions. The material properties of these models are summarized in Table 5.2 and their stress-strain functions are depicted in Figure 5.3 and 5.4. The relative stress-strain relationships are shown in Figure 5.5.
Hyperelastic Models: Two isotropic hyperelastic models were tested. The neo-Hookean model as defined by (3.36) and the 5-parameter Mooney-Rivlin model (3.38). These models were labelled MM_inH and MM_iMR respectively, when fitted to the strain-stress relationship of MM_i for strains $e \in [-0.25, 0.25]$ in the least squares sense. Parameter fitting with ANSYS led to instable materials for the Mooney-Rivlin models. A constrained parameter fitting was developed instead. The employed constraints ensured that the work required for an arbitrary change in deformation is always positive, i.e. $d\sigma/de > 0$ for $e \in [-0.9, 9]$. No convergence problems were experienced for models derived in this way. The parameters of these models are listed in Table 5.3 and their stress-strain relationships are shown in Figure 5.6 to 5.8 in comparison to the original functions. It can be observed that neo-Hookean models approximated linear models well, but were not flexible enough for the non-linear functions. Better fits were achieved using the Mooney-Rivlin models due to the greater number of free parameters.

Transverse Isotropic Models: Cooper ligaments are likely to lead to an increased global stiffness in the posterior-anterior direction of the breast, since they connect the deep facia with the skin. Transverse isotropic models, as described in Section 3.3.1, were employed to test this hypothesis in Test5. For an isotropic volume change behaviour and displacement boundary conditions, these models can be described by three parameters $E_3/E_p$, $\mu_3p$ and $v$. $E_3/E_p$ denotes the ratio of the Young's modulus in posterior-anterior direction ($E_3$) to the Young's modulus in the other two directions ($E_p$). $\mu_3p$ represents the shear modulus in the posterior-anterior direction and $v$ stands for the isotropic Poisson’s ratio. Tests were conducted for $E_3/E_p \in [0.2, 7.0]$ and $\mu_3p \in [0.2, 4.2]$. The material properties of fatty and fibroglandular tissue were assumed to be the same in these tests. The Poisson’s ratio was set to a suitable value as determined during tests of the linear isotropic homogeneous model (MM1) as determined during Test4.

Poisson’s ratio The influence of the Poisson’s ratio was assessed for values 0.499, 0.495, 0.45, 0.4, 0.3, 0.2 and 0.1 for the linear and piecewise linear material models in Test4.
### Table 5.2: Young’s modulus $E$ of fatty and fibroglandular tissue for linear and piecewise-linear material models. The last column lists mean($E_g/E_f$) for strain $e \in [-0.25, 0]$.

<table>
<thead>
<tr>
<th>Name</th>
<th>Young’s modulus $E$ in kPa for</th>
<th>$E_g/E_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fatty tissue ($E_f$)</td>
<td>Fibroglandular tissue ($E_g$)</td>
</tr>
<tr>
<td><strong>Linear</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MM2</td>
<td>1</td>
<td>1.5</td>
</tr>
<tr>
<td>MM3</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>MM4</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>MM5</td>
<td>1</td>
<td>15</td>
</tr>
<tr>
<td>MM6</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td><strong>Non-linear</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM7</td>
<td>718.16$e + 4.46$ if $e &lt; 0.25$</td>
<td>15.1 exp(10.0$e$)</td>
</tr>
<tr>
<td></td>
<td>184 if $e \geq 0.25$</td>
<td></td>
</tr>
<tr>
<td>MM8</td>
<td>18.5</td>
<td>1001.3$e^2 - 272.7e + 86.1$</td>
</tr>
<tr>
<td>MM9</td>
<td>519.7$e^2 + 2.4e + 4.9$</td>
<td>123888.9$e^3 - 11766.7e^2 + 696.9e + 12.1</td>
</tr>
</tbody>
</table>

### Table 5.3: Parameters of strain energy function $\Psi$ for hyperelastic material models as defined in (3.36) and (3.38). The last column states mean($E_g/E_f$) for strain $e \in [-0.25, 0]$.

<table>
<thead>
<tr>
<th>Name</th>
<th>Parameters of strain energy function $\Psi$ in kPa for</th>
<th>$E_g/E_f$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Fatty tissue</td>
<td>Fibroglandular tissue</td>
</tr>
<tr>
<td></td>
<td>$\alpha_1$</td>
<td>$\alpha_1$</td>
</tr>
<tr>
<td><strong>Neo-Hookean</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM1nH</td>
<td>0.1289</td>
<td>0.1289</td>
</tr>
<tr>
<td>MM2nH</td>
<td>0.1289</td>
<td>0.1933</td>
</tr>
<tr>
<td>MM3nH</td>
<td>0.1289</td>
<td>0.6443</td>
</tr>
<tr>
<td>MM4nH</td>
<td>0.1289</td>
<td>1.2887</td>
</tr>
<tr>
<td>MM5nH</td>
<td>0.1289</td>
<td>1.9330</td>
</tr>
<tr>
<td>MM6nH</td>
<td>0.1289</td>
<td>2.5774</td>
</tr>
<tr>
<td>MM7nH</td>
<td>19.8360</td>
<td>17.3002</td>
</tr>
<tr>
<td>MM8nH</td>
<td>2.4134</td>
<td>11.2613</td>
</tr>
<tr>
<td>MM9nH</td>
<td>3.5714</td>
<td>105.0993</td>
</tr>
<tr>
<td><strong>Mooney-Rivlin</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MM7MR</td>
<td>$\alpha_{10}$ $\alpha_{01}$ $\alpha_{20}$ $\alpha_{11}$ $\alpha_{02}$</td>
<td>$\alpha_{10}$ $\alpha_{01}$ $\alpha_{20}$ $\alpha_{11}$ $\alpha_{02}$</td>
</tr>
<tr>
<td></td>
<td>46.42 -31.77 37.07 1.96 1.51</td>
<td>42.83 -36.54 51.83 7.33 0.52</td>
</tr>
<tr>
<td>MM8MR</td>
<td>5.83 -3.14 0.90 0.64 0.08</td>
<td>26.07 -15.56 8.10 1.71 0.31</td>
</tr>
<tr>
<td>MM9MR</td>
<td>10.00 -7.14 3.12 1.82 2.07</td>
<td>263.15 -231.32 387.50 55.81 -0.02</td>
</tr>
</tbody>
</table>
Figure 5.3: Stress-strain relationships of fatty tissue for MM1-MM9. The right figure shows details for the low stress values.

Figure 5.4: Stress-strain relationships of fibroglandular tissue for MM1-MM9.

Figure 5.5: Relative stress-strain relationships of fibroglandular to fatty tissue for MM1-MM9.
5.2 Materials and Methods

Figure 5.6: Stress-strain relationships of fatty tissue for hyperelastic models.

Figure 5.7: Stress-strain relationships of fibroglandular tissue for hyperelastic models.

Figure 5.8: Relative stress-strain relationships of fibroglandular to fatty tissue for hyperelastic models.
5.2 Materials and Methods

Table 5.4: Configurations of three displacement boundary conditions (BC1, BC2 and BC3).

The surface nodes of the FE model were classified as posterior, medial and lateral as illustrated by the black areas on the left figures. Surface nodes were either free to relax (free), displaced using the result of the 3D non-rigid registration (nreg) or fixed to remain at their initial position (fixed). BC3 contained additional nodes on the lateral and medial side towards the posterior side to compensate for the completely unconstrained posterior side.

5.2.3.3 Boundary Conditions

The influence of boundary conditions was assessed in Test4 for three different setups, called BC1, BC2 and BC3. Displacements were prescribed for the surface nodes on the lateral, medial and posterior side of the breast as illustrated in Table 5.4. The posterior nodes were defined to be the subset of nodes which were visible when viewing the opaque triangulated breast surface from the posterior side using The Visualization Toolkit [Schroeder et al., 1998]. Surface nodes which were within 40mm lateral distance of the most medial node while not being posterior of the compression plates were selected to belong to the medial side and likewise for the surface nodes within 40mm of the lateral side. BC1 and BC2 constrained 51% (54%) of the surface nodes for volunteer one (two), while BC3 constrained 24% (46%).

BC1: Displacements derived from the registration were applied to the posterior, medial and lateral nodes. No boundary conditions were imposed on any of the remaining nodes.

BC2: Assuming that the pectoral muscle and the medial side of the breast had not moved, the posterior and medial nodes were set to have zero displacements. Displacements from the registration were applied to the lateral nodes.

BC3: This boundary condition is the same as BC2, apart from not constraining the posterior side and increasing the number of nodes for the lateral and medial side.

The volume change introduced by a boundary condition was estimated from the volume of the original and the warped surface triangulation. For BC1, the triangulation was warped...
by the 3D non-rigid image registration, because all displacements prescribed by BC1 were defined by the outcome of the image registration. This resulted in a volume change of 0.1% (-5.9%) for volunteer one (two) and hence we concluded that the deformation of the breast of volunteer one did not change volume. For BC2 and BC3, the triangulation was warped by the FEM model which provided the best mean displacement error for the boundary condition, resulting in an estimated volume change of -6.4% (-6.5%) for BC2 and -1.0% (-9.45%) for BC3.

5.2.3.4 FEM Model Solution

All FEM model solutions were obtained using a static or steady state analysis, i.e. the models were assumed to have relaxed to their lowest energy solution. The general approach was to use the preconditioned conjugate gradient solver; to employ ANSYS’s automatic load step adjustment with initially 1 step (linear models) or 4 steps (non-linear models) and maximally 500 steps; and to use an infinitesimal deformation formulation. For volunteer one, this strategy was tested against the direct solver (Test1), and a finite deformation formulation (Test3), i.e. taking account of the geometric non-linearity.

For pure displacement approaches, nearly incompressible material can lead to predicted displacements which are much smaller than they should be, i.e. locking, or to no convergence, see Section 3.4.4. The mixed u-p formulation, where pressure is another solution variable, was employed to overcome these problems for the finite deformation formulation. Piecewise-linear functions are, however, not supported by the mixed u-p formulation by ANSYS. Non-linear models were instead approximated by hyperelastic models.

The ANSYS solution provides displacements at all node positions. A continuous displacement field was derived from this by interpolation with the quadratic shape function for 10-noded tetrahedral elements as defined in Section 3.4.1.

5.2.4 Landmark Selection

The accuracy of each model was assessed by manually identifying for each breast 12 corresponding point landmarks in the set of images.

Reference landmarks were selected in the image of the deformed breast. Landmarks were generally placed at bifurcations, centres of small tissue regions or tips of spiculated branches, as illustrated in Figure 5.9 and 5.10. Two landmarks were usually picked in every 10th slice for volunteer one and in every 7th slice for volunteer two. Figure 5.11 shows the distribution of the landmarks.

The corresponding landmarks were then identified in the image showing the undeformed
5.2 Materials and Methods

breast. This was repeated two more times by a single observer on different days. The average position of these three measurements was assigned to each landmark.

Figure 5.11 shows the displacement pattern of the landmarks for both volunteers. These differ mainly in inferior to superior direction. This was caused by the inferior part (volunteer one) or superior part (volunteer two) of the undeformed breast being closer to the lateral plate.

5.2.5 Error Measures

The intra-observer variability of selecting a landmark was measured by the mean Euclidean distance of the three measurements to their average position. The table on the right lists the minimum, mean, 90th percentile and maximum results for the 12 landmarks of each volunteer. The 24 landmarks were selected with a mean (90th percentile) intra-observer variability of 0.24mm (0.50mm).

The performance of a model was quantified by the Euclidean distance between the position of the reference landmark and the predicted position of the corresponding landmark by the model. This distance is called the displacement error.

For a given volunteer and boundary condition, any model that produced a mean (maximum) displacement error that exceeded the overall best results by less than the mean (90th percentile) intra-observer variability of 0.24mm (0.50mm) was also classified as best. Statistical significance was not assessed because of the small number of volunteers.

The sensitivity of the displacement errors was assessed with respect to 4 factors, namely the mean Young’s moduli ratio of fibroglandular and fatty tissue, the Poisson’s ratio, the boundary condition and the addition of skin. The analysis was based on the model-free sensitivity index $S_i$, which measures the average reduction in variance of the displacement error for sample quantity $Y$ when fixing individual factors $F_i$. $S_i$ is given by $1 - (\frac{1}{N} \sum_{j=1}^{N} \text{var}(Y|F_i = f_j))/\text{var}(Y)$ where $N$ is the number of possible values for factor $F_i$ and $\text{var}(Y|F_i = f_j)$ is the variance of $Y$ after fixing factor $F_i$ to value $f_j$ [Saltelli et al., 2004].
Figure 5.9: Examples of corresponding landmarks for volunteer one. The images show orthogonal 2D slices of the undeformed breast (left) and the deformed breast (right). The landmarks are located at the centre of each circle.
Figure 5.10: Examples of corresponding landmarks for volunteer two. The images show orthogonal 2D slices of the undeformed breast (left) and the deformed breast (right). The landmarks are located at the centre of each circle.
Figure 5.11: Distribution and displacement of landmarks in the deformed breast. The images show the orthogonal mean intensity projections of the breast segmentation. The projected landmark positions are marked by o. Lines show the projected displacements of the landmarks. The minimum, mean and maximum displacement of all 12 landmarks are listed for each direction ($u_1$: lateral to medial, $u_2$: posterior to anterior, $u_3$: inferior to superior). $u$ denotes the length of the displacements.

5.2.6 Overview of Tests

Five tests were performed in this study. The purpose of these tests are described next and an overview of their configuration is given in Table 5.5.

**Test1 - Direct Solvers:** Results obtained with the preconditioned conjugate gradient solver were compared to results obtained by two direct solvers, namely the frontal and the sparse direct solver.

**Test2 - Mesh Density:** This test assessed whether doubling the number of mesh elements provided different results.

**Test3 - Finite Deformation Formulation:** A finite and an infinitesimal deformation formulation were compared in Test3. Incompressible hyperelastic models and a mixed $u$-$p$ approach were used for the finite deformation formulation.

**Test4 - Linear and Piecewise-Linear Models, Skin, Poisson’s Ratios, Boundary Conditions:** This is the main test. Six linear and three piecewise-linear models with and without skin were compared for three boundary conditions and at least six Poisson’s ratios.
Test5 - Transverse Isotropic Models: This test assessed if transverse isotropic models improve the results of the isotropic homogeneous model without skin (MM1).

5.2.7 Summary

This section explained the design of the biomechanical breast models, the parameters that were tested and the evaluation strategy. The results obtained for Test1 to Test5 are described in the next section.

<table>
<thead>
<tr>
<th>Section</th>
<th>Configuration</th>
<th>Test1</th>
<th>Test2</th>
<th>Test3</th>
<th>Test4</th>
<th>Test5</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2.1 Volunteer</td>
<td>One</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2.3.1 Mesh</td>
<td>GM1</td>
<td>x</td>
<td>A</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>GM2</td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2.3.2 Material model</td>
<td>Linear</td>
<td>x</td>
<td>x</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>Piecewise-linear</td>
<td>x</td>
<td>x</td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hyperelastic</td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Transverse isotropic</td>
<td></td>
<td></td>
<td></td>
<td>B</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Without skin</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>C</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>With skin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>5.2.3.2 Poisson’s ratio</td>
<td>0.499</td>
<td></td>
<td></td>
<td>A</td>
<td>x</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.495</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.2</td>
<td></td>
<td></td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.499-0.1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>E</td>
</tr>
<tr>
<td>5.2.3.3 Boundary condition</td>
<td>One</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>F</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Two</td>
<td></td>
<td></td>
<td></td>
<td>G</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Three</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>H</td>
</tr>
<tr>
<td>5.2.3.4 FEM solver</td>
<td>PCG</td>
<td>A</td>
<td>x</td>
<td>A</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Frontal</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sparse direct</td>
<td>C</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Mixed u-p</td>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.2.3.4 Deformation formulation</td>
<td>Infinitesimal</td>
<td>x</td>
<td>x</td>
<td>A</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td></td>
<td>Finite</td>
<td></td>
<td></td>
<td>B</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.5: Overview of test configurations. The entries 'A' to 'H' mark for each test the aspects that were compared. Unchanged factors are denoted by 'x'.
5.3 Results

This study assessed the influence of different elasticity values and Poisson’s ratios (Section 5.2.3.2), boundary conditions (Table 5.4), finite element solvers (Section 5.2.3.4) and mesh resolutions (Section 5.2.3.1) on the accuracy with which biomechanical breast models can predict the displacements of internal breast structures. Firstly, the range of residual misalignment of the tested models is illustrated by visual example results. Thereafter, the quantitative results of the five tests are presented.

5.3.1 Visualisation of Results

Figure 5.12(a) and 5.13(a) show 2D orthogonal example slices of the MR breast images before and after compressing the breast between two plates for volunteer one and two respectively. The misalignment due to the compression is clearly visible on the difference images. Motion artifacts in the difference images were greatly reduced after 3D non-rigid registration (Figure 5.12(b) and 5.13(b)) with the mean (maximum) displacement error dropping from 6.41mm (12.98mm) to 0.92mm (1.67mm) for volunteer one and from 6.80mm (11.85mm) to 1.03mm (2.74mm) for volunteer two. Less artifact reduction was achieved by the biomechanical breast models, with the mean displacement error ranging from 1.85mm to 4.58mm for volunteer one (Figures 5.12(c-h)) and from 2.12mm to 4.01mm for volunteer two (Figures 5.13(c-h)).

5.3.2 Test1: Direct Solvers

Employing the direct solvers (frontal or sparse direct) instead of an iterative solver changed the mean (maximum) displacement errors of the material models MM1 to MM9 by less than 0.0007mm (0.0012mm), see Table 5.6. Results of the frontal and sparse direct solver were identical for the stated precision.

5.3.3 Test2: Mesh Density

A denser mesh (GM2) changed the mean (maximum) displacement error on average by 0.003mm (0.004mm) for volunteer one and by -0.088mm (-0.059mm) for volunteer two. The mean (maximum) displacement error of the individual models changed by less than 0.097mm (0.196mm) for both volunteers. (Table 5.6).
Figure 5.12: Axial, sagittal and coronal 2D example slices of volunteer one showing (a) original images with (top) compressed breast, (middle) uncompressed breast, (bottom) difference image of uncompressed and compressed breast; (b) difference image after 3D non-rigid registration; (c-h) difference image of the FEM deformed image and the original image of the compressed breast for results with (c,e,g) smallest mean displacement error and (d,f,h) largest mean displacement error for BC1-3. The mean displacement error DE are listed for each case.
Figure 5.13: Axial, sagittal and coronal 2D example slices of volunteer two (a) of the original image with (top) compressed breast, (middle) uncompressed breast, (bottom) difference image of uncompressed and compressed breast; (b) difference image after 3D non-rigid registration; (c-h) difference image of the FEM deformed image and the original image of the compressed breast for results with (c,e,g) smallest mean displacement error and (d,f,h) largest mean displacement error for BC1-3. The mean displacement error DE are listed for each case.
Table 5.6: Change of displacement errors due to the use of different FEM solvers and geometric models for volunteer one, boundary condition BC1 and Poisson’s ratio $v = 0.495$ ($v = 0.2$) for volunteer one (two). The first column (GM1, PCG) lists the mean and maximum displacement errors for the geometrical model GM1 employing the preconditioned conjugate gradient (PCG) solver. The columns labelled with $\Delta$mean ($\Delta$max.) state the change of the mean (maximum) displacement error due to the use of a different solver or geometrical model. The mean value for all 9 material models is given in the last row.
5.3.4 Test3: Finite Deformation Formulation (Volunteer 1, BC1 only)

Finite deformation solutions were obtained for incompressible hyperelastic models using a mixed u-p solution strategy for the volume preserving deformation, i.e. volunteer one and boundary condition BC1. These were compared to linear and piecewise-linear models that used an infinitesimal deformation formulation and a Poisson's ratio of 0.499.

None of the finite deformation formulations improved the result of their infinitesimal counterpart, see Table 5.7. Neo-Hookean models increased the mean (maximum) displacement errors on average by 0.07mm (0.15mm) when compared to the linear models and by 0.28mm (0.34mm) for MM7-MM9. Approximating the non-linear material models (MM7-MM9) by 5-parameter Mooney-Rivlin models increased the mean (maximum) displacement error by 0.14mm (0.25mm) on average. The volume changes introduced by the hyperelastic models (0.0% to 0.1%) resembled the volume change of the 3D image registration. The linear and piecewise-linear models reduced the mesh volume by 0.7%.

<table>
<thead>
<tr>
<th>Finite deformation, mixed u-p formulation</th>
<th>Infinitesimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neo-Hookean model</td>
<td>Mooney-Rivlin model</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td></td>
<td>(mm)</td>
</tr>
<tr>
<td>MM1nH</td>
<td>2.00</td>
</tr>
<tr>
<td>MM2nH</td>
<td>2.02</td>
</tr>
<tr>
<td>MM3nH</td>
<td>2.22</td>
</tr>
<tr>
<td>MM4nH</td>
<td>2.43</td>
</tr>
<tr>
<td>MM5nH</td>
<td>2.60</td>
</tr>
<tr>
<td>MM6nH</td>
<td>2.72</td>
</tr>
<tr>
<td>Mean (1-6)</td>
<td>2.33</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>MM7nH</td>
<td>2.00</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>MM8nH</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (7-9)</td>
<td>2.36</td>
</tr>
</tbody>
</table>

Table 5.7: Comparison of the finite deformation formulation results of neo-Hookean models (MM1nH-MM9nH) and Mooney-Rivlin models (MM7MR-MM9MR) with the infinitesimal deformation formulation results of linear or piecewise-linear models (MM1-MM9) for volunteer one and boundary condition BC1.
The mean (maximum) displacement error changed by more than 0.24mm (0.50mm) only when approximating MM9 by a neo-Hookean model. This worse performance is likely to have been caused by the poor approximation of the stress-strain relationship as is apparent in Figures 5.6-5.8. The neo-Hookean models matched the linear models very well, but still did not improve the results. It must therefore be concluded, that incompressible hyperelastic models and the use of a finite deformation formulation results in a similar accuracy to linear and piecewise-linear models with a Poisson’s ratio of 0.499 and an infinitesimal deformation formulation for a volume preserving deformation.

5.3.5 Test4: Linear and Piecewise-Linear Models, Skin, Poisson’s Ratios, Boundary Conditions

The effects of the average elastic ratio of fibroglandular and fatty tissue (mean($E_g/E_f$)) and the Poisson’s ratio $v$ on the resulting displacement error are illustrated for models without skin by contour plots in Figure 5.14 to 5.19. Table 5.8 provides an overview of the main results.

- For volunteer one and boundary condition BC1, the best results were achieved by material models where fibroglandular tissue was at most 4 times stiffer than fatty tissue and by MM3S while keeping Poisson’s ratios high ($v \in \{0.499, 0.495\}$), see Figure 5.14. These models reduced the mean (maximum) displacement error to 1.96mm (3.43mm). Assuming that no motion had occurred at the deep and medial side (BC2) caused high maximum errors, see Figure 5.15. Best results for BC2 required low Poisson’s ratios ($v \in \{0.1, 0.2, 0.3\}$), due to the introduction of volume changes. Keeping the posterior side unconstrained (BC3) was better than assuming no motion at the posterior side (BC2) but was worse than BC1. Best models for BC3 had low Young’s moduli ratios and high Poisson’s ratios, see Figure 5.16. Very similar contour plots were achieved when adding a skin with a Young’s modulus of 10kPa, with the mean (maximum) displacement error of any model changing by less than 0.21mm (0.36mm). The mean and maximum displacement error was most sensitive to the boundary condition, see Table 5.9. The most important factor after fixing the boundary condition was either the Poisson’s ratio (for BC1 apart from the mean error of MM1-MM6, for BC2 and for the maximum error of MM7-MM9 for BC3) or the Young’s moduli ratio. Thin plate spline interpolation of the prescribed surface displacements provided only for BC1 a

\footnote{Best results are defined as results within the mean (90th percentile) intra-observer variability of 0.24mm (50mm) from the lowest mean (maximum) displacement error, see Section 5.2.5.}
5.3 Results

mean and a maximum displacement error that was similar to the best FE models, see Table 5.8.

- For volunteer two, the non-rigid registration introduced a volume change of $-5.9\%$. Substantially worse results were therefore obtained at high Poisson’s ratios, see Figure 5.17 and Table 5.8. Best results for BC1 were achieved by all models but MM6, MM9, MM9S while keeping Poisson’s ratios low ($v \in \{0.1, 0.2, 0.3\}$). Assuming no motion at the deep and medial side (BC2) introduced lower errors than for volunteer one (Table 5.8). Not constraining the posterior side (BC3) provided similar results as BC2. Best results for BC3 were achieved by linear models with skin and a low Poisson’s ratio. Adding skin had a bigger impact on the results than for volunteer one, with the mean (maximum) displacement error changing on average by -0.09mm (-0.42mm) and up to 0.31mm (2.21mm). The mean and maximum displacement error was most sensitive to the Poisson’s ratio, see Table 5.9. The most important factor after fixing the Poisson’s ratio was the boundary condition apart from one case (mean error of MM7-MM9 for $v=0.1$) where the results were more sensitive to the Young’s moduli ratio. Thin plate spline interpolation of the boundary constraints (Table 5.8) provided worse results than the best FE models.

In conclusion, nine material models achieved for BC1 and appropriate Poisson’s ratios best results for both volunteers, namely all material models where fibroglandular tissue was at most 4 times stiffer than fatty tissue and MM3S. Their mean (maximum) displacement error was on average 2.11mm (3.37mm). Less accurate or less constrained boundary conditions (BC2, BC3) led to higher errors. Keeping the posterior side unconstrained (BC3) was better for volunteer one than assuming no motion at this side (BC2). The boundary condition or the Poisson’s ratio had the highest influence on the displacement error.
5.3 Results

Linear models MM1-MM6

Non-linear models MM7-MM9

Figure 5.14: Results for volunteer one, boundary condition BC1 and material models without skin. The mean and maximum displacement errors are displayed as contour plots with respect to the mean($E_g/E_f$) of 9 material models (x-axis, see Table 5.2) and 7 Poisson’s ratios (y-axis: 0.499, 0.495, 0.45, 0.4, 0.3, 0.2, 0.1). The contour interval is 0.1mm. Only the 5 lowest contours of each graph are labelled. The largest value of each graph is shown by the boxed number. Result within the mean (90th percentile) intra-observer variability of 0.24mm (0.50mm) from the smallest mean (maximum) displacement error of 1.88mm (3.33mm) are marked by stars. Dots represent the remaining experiments.

Figure 5.15: Results for volunteer one, boundary condition BC2 and material models without skin. The lowest mean (maximum) displacement error for BC2 was 2.75mm (7.12mm).
5.3 Results

**Linear models**

MM1-MM6

**Non-linear models**

MM7-MM9

Figure 5.16: Results for volunteer one, boundary condition BC3 and material models without skin. The lowest mean (maximum) displacement error for BC3 was 2.25mm (4.76mm).

Figure 5.17: Results for volunteer two, boundary condition BC1 and material models without skin. The smallest mean (maximum) displacement error for BC1 was 2.12mm (3.15mm).
5.3 Results

Linear models MM1-MM6

Non-linear models MM7-MM9

Figure 5.18: Results for volunteer two, boundary condition BC2 and material models without skin. The smallest mean (maximum) displacement error for BC2 was 2.36mm (3.85mm).

Figure 5.19: Results for volunteer two, boundary condition BC3 and material models without skin. The smallest mean (maximum) displacement error for BC2 was 2.52mm (3.80mm).
### 5.3 Results

<table>
<thead>
<tr>
<th>Smallest error (mm)</th>
<th>Best FE models</th>
<th>TPS interp.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Material model</td>
<td>Poisson’s ratio</td>
</tr>
<tr>
<td>Volunteer one</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BC1</td>
<td>1.88 3.33</td>
<td>*mean(E_g/E_f) &lt; 5, MM3S  0.495, 0.499</td>
</tr>
<tr>
<td>BC2</td>
<td>2.75 7.12</td>
<td>*mean(E_g/E_f) &lt; 4, MM9(S)  0.1-0.3</td>
</tr>
<tr>
<td>BC3</td>
<td>2.25 4.76</td>
<td>*mean(E_g/E_f) &lt; 4, MM8S  0.495, 0.499</td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td></td>
</tr>
<tr>
<td></td>
<td>mean</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.8: Configuration and mean results of the best FE models for Test4, in comparison to the smallest displacement errors of the FE models and the results from a thin plate spline (TPS) interpolation. Interpolation results within 0.24mm (0.50mm) from the smallest mean (maximum) displacement error are marked by stars.

#### Sensitivity Analysis

<table>
<thead>
<tr>
<th>Fixed factor</th>
<th>Sensitivity index $S_i$ in %</th>
<th>Volunteer one</th>
<th>Volunteer two</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MM1-MM6</td>
<td>MM7-MM9</td>
<td>MM1-MM6</td>
</tr>
<tr>
<td>Fixed factor</td>
<td>Mean</td>
<td>Max.</td>
<td>Mean</td>
</tr>
<tr>
<td>Young’s moduli ratio $E_g/E_f$</td>
<td>15</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Poisson’s ratio $v$</td>
<td>5</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Boundary condition</td>
<td>52</td>
<td>92</td>
<td>59</td>
</tr>
<tr>
<td>Skin</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5.9: Sensitivity analysis of the displacement error. The sensitivity index $S_i$ measures the average variance reduction of the mean or maximum displacement error when fixing the listed factor. Influential factors are indicated by large positive $S_i$ values.

### 5.3.6 Test5: Transverse Isotropic Models

This test assessed whether the results from material model MM1 can be improved by transverse isotropic models.

- For volunteer one, transverse isotropic materials with an isotropic Poisson’s ratio of $v = 0.499$ achieved best mean results for a wide range of Young’s moduli ratios $E_3/E_p$ and shear moduli $\mu_3p$ as shown in Figure 5.20. Transverse isotropic models reduced the mean (maximum) displacement error of the isotropic homogeneous model by at most 0.04mm (0.43mm).
For volunteer two, no improvement was achieved by transverse isotropic models with $v = 0.2$, see Figure 5.21. The mean displacement error was more influenced by the shear modulus than the Young's moduli ratio.

Transverse isotropic models with suitable parameters did not substantially improve the model accuracy.

Figure 5.20: Results for volunteer one, boundary condition BC1, Poisson’s ratio $v = 0.499$ and transverse isotropic material models. The mean and maximum displacement errors are displayed as contour plots with respect to the ratio of the stiffness in posterior-anterior direction ($E_3$) to the stiffness in the other directions ($E_p$) (from 0.6 to 7.0 on x-axis) and the shear modulus $\mu_{3p}$ (from 0.2 to 4.2 on y-axis). The smallest mean (maximum) displacement error was 1.84mm (2.91mm).

Figure 5.21: Results for volunteer two, BC1, $v = 0.2$ and transverse isotropic material models. The smallest mean (maximum) displacement error was 2.40mm (3.43mm).
5.4 Summary

The first aim of the presented study was to determine a suitable FEM configuration which can produce plausible deformations during DCE MR mammography acquisition. For this purpose, any model which achieved, for accurate boundary conditions, a mean (maximum) displacement error that exceeded the best result by less than 0.24mm (0.50mm) for both volunteers was regarded as suitable. The second aim was to investigate what modelling aspects are most important.

Biomechanical breast models were able to predict the displacements of internal breast structures (i.e. twelve corresponding anatomical landmarks) for accurate boundary conditions (BC1), appropriate Poisson’s ratios and suitable elastic properties (fibroglandular tissue is at most 4 times stiffer than fatty tissue) to a mean accuracy of 2.0mm (volunteer one) and 2.2mm (volunteer two) for a deformation which introduced a mean displacement of 6.4mm and 6.8mm, respectively. These models reduced on average the maximum displacement error from 13.0mm to 3.4mm for volunteer one and from 11.8mm to 3.3mm for volunteer two. These suitable elastic properties are within the range of values reported by in-vivo elastography studies [Lawrence et al., 1998a; Lawrence et al., 1999; McKnight et al., 2001; van Houten et al., 2003].

The mean displacement errors of 378 models (9 Young’s moduli ratios \times 7 Poisson’s ratios \times 3 boundary conditions \times 2 skin properties) were most sensitive to the boundary condition for volunteer one ($S_i \geq 52\%$) and to the Poisson’s ratio for volunteer two ($S_i \geq 39\%$). The sensitivity was measured by $S_i$, which denotes the average variance reduction of the errors when fixing a factors. Lower Poisson’s ratios improved the prediction when volume changes needed to be modelled. This is the first study that assessed the influence of the Poisson’s ratio and the Young’s moduli ratio between fibroglandular and fatty tissue on the prediction accuracy in a systematic way.

Assuming no motion at the posterior and medial nodes, as often done when modelling breast compressions [Azar et al., 2000; Azar et al., 2002; Samani et al., 2001b; Yin et al., 2004; Pathmanathan et al., 2004; Ruiter et al., 2003; Ruiter, 2003], led to worse results for volunteer one than keeping the posterior nodes unconstrained.

The robustness of the results to the elastic properties were probably caused by modelling only a few tissue types, and by the FEMs being greatly constrained by surface displacements and hence acting more like interpolants. A similar robustness was reported in [Ruiter, 2003].

Employing direct solvers or a finer mesh of improved quality did not influence the ac-
accuracy. Using hyperelastic material models and a finite deformation formulation did not improve the accuracy. That these computationally more expensive models did not improve the accuracy is an encouraging results for applications where fast models are required.

Transverse isotropic models with suitable parameters provided similar results as a homogeneous, isotropic model. This was the first time that transverse isotropic material models were tested for modelling breast deformations.

The results of this study compare favourably with [Azar et al., 2000] and [Ruiter, 2003], most likely because of improved boundary conditions and much finer mesh resolution. The 3D image registration achieved mean errors of about 1.0mm using just the information provided by the images and a mechanical unconstrained transformation model. Whether such an accuracy could be obtained prospectively with a patient specific, heterogeneous biomechanical model and matching boundary conditions, remains to be seen.

The aim of this study was to determine a good model configuration when employing displacement boundary conditions. The parameter assessment can therefore be restricted to the Young’s moduli ratios of the different tissue types, and the forces applied in the experiments were not measured. Hence it is unknown whether the 6% volume change observed for volunteer two was caused by the applied forces exceeding the blood pressure values. During DCE MR mammography acquisition, however, no significant change in external force is applied.

In conclusion, firstly all material models where fibroglandular tissue was no greater than 4 times stiffer than fatty tissue provided the best results for both volunteer and can be recommended in conjunction with a high Poisson’s ratio for the simulation of plausible breast deformations during DCE MR mammography. Secondly, the accuracy of the model was more affected by the choice of boundary condition or Poisson’s ratio than by the choice of elastic material properties. Further studies are required to reconfirm the importance of these modelling factors for other applications.
Chapter 6

Validation of DCE MR Mammography Registration

This chapter describes a validation study of a registration algorithm for aligning the image sequence of DCE MR breast images acquired during one visit. The study was based on simulating plausible breast deformations using a suitable biomechanical breast model as determined in Chapter 5. The content of this chapter is based on and extends [Tanner et al., 2002b; Schnabel et al., 2003]. The author developed the biomechanical models, conducted the tests and analysed the results described in this chapter.

6.1 Introduction

As discussed in Chapter 1.5, correct alignment of intra-visit DCE MR mammograms supports the detection and diagnosis of breast lesions. Many registration methods have been proposed for this task, but evaluation of their performance was insufficient, see Chapter 2.

To improve this situation, this study employed a novel validation method to estimate the accuracy of intra-visit DCE MR mammography registration. The validation method is based on applying the registration to misaligned images, generated from plausible deformations simulated by biomechanical models using finite element methods. The configuration of the finite element model was chosen using measurements of breast deformations on two volunteers, see Chapter 5. Gold standard breast deformations were computed for ten patients with twelve deformation scenarios each. These deformations were applied to the post-contrast image to simulate patient motion occurring between pre- and post-contrast image acquisition. The original pre-contrast images were registered to the corresponding deformed post-contrast images. The most accurate registration protocols for DCE MR mammography were then determined using the misaligned image pairs of five patients (training set). Subsequently, the expected registration accuracy for these strategies was estimated from an independent test set consisting of images from the other five patients. Additionally, the sensitivity of the registration results to various factors was evaluated and the results were tested for statistical significance.

The remainder of this chapter is structured as follows. Section 6.2 describes the design of the gold standard deformations and the evaluation strategy. The results are presented in
6.2 Material and Methods

This section describes the gold standard deformations and the evaluation strategy.

6.2.1 Datasets

Ten patients, where no motion between image acquisitions was visible, were selected by the author from 147 symptomatic patients. Sixty-seven of these were patients from the Guy’s and St. Thomas’ Hospital Trust. The remaining 80 patients were enrolled in the UK MR breast screening study (MARIBS) [Brown et al., 2000]. The selected images were acquired with a 3D gradient echo sequence on various 1.5T MR systems with TR=20ms, TE=5.2ms, flip angle = 45° (patients 1-4) or TR = 12ms, TE = 5ms, flip angle = 35° (patients 5-10). The voxel dimensions are 1.37×1.37×4.2mm (patients 1,3,4), 1.48×1.48×4.2mm (patient 2) and 1.33×1.33×2.5mm (patients 5-10, see [Brown et al., 2000]). The slice orientation is axial (patients 1-4) or coronal (patients 5-10). Example slices of the intensity difference of the unregistered images are shown in Figure 6.5(a) and 6.6(a) on page 191 and 192, where images of patients 5-10 were reformatted to have an axial slice direction for better visual comparison.

6.2.2 Biomechanical Breast Models

The evaluation of the registration method is based on creating gold standard transformations by simulating plausible breast deformations using biomechanical breast models based on finite element methods.

Meshes were created for the 10 selected image pairs in a similar way as described in Section 5.2.3.1. The images were segmented into fat, fibroglandular tissue and enhancing lesion. The outside surface of the breasts and of the enhancing lesions were then triangulated at a scale of 8.4mm to 12.5mm for the breasts and of 4.2mm to 7.5mm for the enhancing lesions, such that the limit of available elements was not exceeded during the meshing stage. The surfaces were meshed into 10-noded tetrahedral elements via ANSYS, resulting in 9.5×10^4 nodes and 6.8×10^4 elements on average. The individual tissue types were modelled as linear, isotropic materials with a Young’s modulus of 1kPa, 1.5kPa and 3.6kPa for fat, fibroglandular tissue and the enhancing lesion, respectively. This was in agreement with the results described in Chapter 5. A Poisson’s ratio of 0.495 was chosen to enforce approximate incompressibility of the tissue.
6.2 Material and Methods

6.2.3 Simulation of Breast Deformations

Twelve deformations were generated. Regional displacement simulated a uniform displacement on one side of the breast; point puncture imitated a deformation during biopsy; one-sided contact simulated a deformation of the breast when pushed against the breast coil; two-sided contact imitated the fixation of the breast between two plates; tension simulated the deformation caused by the contraction of the pectoral muscle; and relaxation imitated the transformation from a contracted to a resting pectoral muscle.

The most frequently observed deformations during MR-mammography are tension and relaxation. The surface displacements of a typical tension and a typical relaxation case were studied. These were imitated by displacements in the posterior and anterior direction with distorted ellipsoidal contours in the axial plane as illustrated in Figure 6.1. Relaxation was simulated as a slender and more distorted shape than tension.

Two magnitudes of displacement were simulated for each deformation scenario. 'Large' simulated the larger deformations (excluding global patient movements) observed during a normal DCE MR mammography session by inducing displacements up to 10mm magnitude. 'Common' imitated the more frequently observed deformation magnitudes with displacements up to 5mm, see Figure 6.2.

For these boundary conditions, the biomechanical breast models were solved using ANSYS [ANSYS Inc., 2000]. A continuous displacement field within the FEM mesh was produced by quadratic shape interpolation of the 10-noded tetrahedral elements as described in Section 3.4.1. This field was applied to the post-contrast image to simulate deformations between pre- and post-contrast images. Image intensities were interpolated using a truncated sinc interpolation kernel (Hanning, radius 6) [Hamming, 1989]. Locations outside the FEM mesh had to be masked out for any further processing since no deformation information is available at these locations. The pre-contrast images were registered to the gold standard deformed post-contrast images to avoid any further interpolation of the latter. Example slices from the difference images after applying these deformations are shown in Figure 6.5b and 6.6b.

6.2.4 Registration

The non-rigid registration tested in this study was based on the multi-resolution free-form deformation (FFD) approach employing B-spline subdivision as described in [Rueckert et al., 1999b] and reviewed in Chapter 2, see page 80. The volume preserving regularization term $C_{\text{volume}}(T)$ suggested by [Rohlfing and Maurer, 2001] was employed. $C_{\text{volume}}(T)$ is
Figure 6.1: Illustration of breast surface displacement due to tension of the pectoral muscle for two right breasts. a) Observed displacement magnitude in the posterior direction with illustration of the pectoralis-major location (black: 7mm, white: -2mm). b) Shape of simulated deformation contours, observed in a). c) Displacements in posterior direction of similar surface displacements applied to a different breast.

Figure 6.2: Histogram of the maximum breast surface displacements after 10mm FFD registration without regularization [Rueckert et al., 1999b] for 31 symptomatic patients. The mean of all patients was 7.60mm.

defined in equation (2.21) on page 61. Optimization was based on the registration cost function [Rohlfing and Maurer, 2001]

\[
C(T) = -(1 - \mu) \, C_{\text{similarity}}(A, B^T) + \mu \, C_{\text{volume}}(T),
\]

where \(C_{\text{similarity}}\) denotes the image similarity measure and \(\mu\) refers to the weight of the volume preserving regularization term which is balancing the two objectives of the cost function.

The registration strategies listed in Table 7.1 were assessed. The registrations were performed over the image region of either the right breast (patients 2, 4, 5, 6 and 9) or the left breast (patients 1, 3, 7, 8 and 10) depending on the location of the lesion, with an average region of interest of \(5.15 \times 10^6\) mm\(^3\) or 0.98 million voxels.
6.2 Material and Methods

### Abbreviation Description of tested registration configurations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description of tested registration configurations</th>
</tr>
</thead>
<tbody>
<tr>
<td>rigid</td>
<td>Rigid registration with 6 degrees of freedom (translation and rotation) and a weight value of $\mu = 0$ in equation (6.1).</td>
</tr>
<tr>
<td>affine</td>
<td>Affine registration with 12 degrees of freedom (translation, rotation, scaling and shearing) and a weight value of $\mu = 0$ in equation (6.1).</td>
</tr>
<tr>
<td>sD$\mu$ W</td>
<td>Single-resolution FFD registration after rigid registration with a control point spacing of $D$mm and a weight value of $\mu = W$ in equation (6.1). Tested for 44 configurations: $D \in {10, 20, 40, 60}$ and $W \in {0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95}$.</td>
</tr>
<tr>
<td>sDc</td>
<td>Single-resolution FFD registration without constraining for volume preservation ($sD\mu 0$) followed by imposing a local rigid transformation in the region of the enhancing lesion using coupled control points, see [Tanner et al., 2000]. Tested for 4 configurations: $D \in {10, 20, 40, 60}$.</td>
</tr>
<tr>
<td>mD$\mu$ W</td>
<td>Multi-resolution FFD registration after rigid registration with an initial control point spacing of 20mm followed by FFD registration(s) with control point spacings of half the previous size up to $D$mm and a weight value $\mu = W$. Tested for 22 configurations: $D \in {5, 10}$ and $W \in {0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95}$.</td>
</tr>
<tr>
<td>mDc</td>
<td>Multi-resolution FFD registration without constraining for volume preservation ($mD\mu 0$) followed by imposing a local rigid transformation within the region of the enhancing lesion using coupled control points, see [Tanner et al., 2000]. Tested for 2 configurations: $D \in {5, 10}$.</td>
</tr>
</tbody>
</table>

Table 6.1: Overview of the 74 registration configurations tested in this validation study.

#### 6.2.5 Error Quantification

A key aspect of the validation method is that the accuracy of the registration can be quantified with respect to the gold standard displacements at each voxel position. The degree of alignment between two corresponding points after registration is described by the Target Registration Error (TRE) [Fitzpatrick, 2001]. Traditionally, TRE is calculated at anatomical landmarks. In the case of FEM-simulated deformations, the correspondence can be estimated at every position within the FEM mesh. Thus, a more evenly sampled error was calculated by computing TRE at all voxel positions $x$ in a region $R$:

$$TRE(x) = ||T_{1F} \circ T_{F2}(x) - x||$$

and

$$TRE_R = \{TRE(x)|x \in R\}, \quad (6.2)$$
where $T_{F2}$ is the FEM-transformation mapping any position in the post-contrast image $I_2$ into the deformed post-contrast image $I_F$. Transformation $T_{1F}$ is obtained from the registration of the pre-contrast image $I_1$ to $I_F$. Equation (6.2) assumes that no motion has occurred between $I_1$ and $I_2$.

TRE distributions were calculated for two regions of interest, namely the whole breast tissue ($TRE_B$) and the enhancing lesion ($TRE_L$) in order to evaluate two registration objectives. One registration aim is to remove motion artifacts within the whole breast tissue to help in the detection of enhancing lesions. The other objective is to accurately align the enhancing lesion to facilitate the lesion’s classification. $TRE_R$, the TRE distribution of a single experiment for region $R$, was described by its mean ($\text{mean}_{TRE_R}$) and its 95th percentile ($95\%_{TRE_R}$). The registration performance for a set of experiments was summarized by computing the mean of all the mean TREs obtained for each experiment ($\text{mean}(\text{mean}_{TRE_B})$, $\text{mean}(\text{mean}_{TRE_L})$) and by calculating the mean of all the 95th percentile TREs obtained for each experiment ($\text{mean}(95\%_{TRE_B})$, $\text{mean}(95\%_{TRE_L})$).

6.2.6 Selecting the Better Registration Strategies

The knowledge of a gold standard not only provides a means to assess the accuracy of a registration method, but also the opportunity to select the better registration parameters. For this purpose, the first 5 image sequences which were accepted as having visually no detectable motion artifacts were assigned to the training set (called patients 1, 2, 5, 6 and 7). Then the remaining 5 selected patients were assigned to the test set (patients 3, 4, 8, 9 and 10). The optimal registration setup was determined as follows.

Firstly, the results for registrations optimized only with respect to one of six voxel-based image similarity measures ($C_{\text{similarity}}$ in equation (6.1), $\mu=0$) were compared for the training set. The tested similarity measures were sum of squared differences defined in equation (2.30) on page 68, cross correlation (2.33), correlation ratio (2.34), joint entropy (2.35), mutual information (2.36) and normalized mutual information (2.37).

Correlation ratio, the only asymmetric measure, was tested for both configurations, i.e. $S_{CR}(B^T|A)$ and $S_{CR}(A|B^T)$. The best similarity measure of this comparison was used for all further registrations.

Secondly, the four TRE sample quantities ($\text{mean}(\text{mean}_{TRE_B})$, $\text{mean}(95\%_{TRE_B})$, $\text{mean}(\text{mean}_{TRE_L})$ and $\text{mean}(95\%_{TRE_L})$) were calculated for the 74 registration configurations described in Table 7.1 from all 60 cases (5 patients $\times$ 12 deformations) of the training set. The 95% confidence intervals of these TRE sample quantities were then approximated by
bootstrapping [Simon, 1997]. For each registration configuration and each TRE sample quantity, this involved the following steps:

1. Sample, with replacement, 60 values from the 60 results (e.g. $95\%_{TRE_B}$).

2. Calculate the mean of these 60 samples.

3. Repeat steps 1.-2. 5000 times.

4. Approximate the 95% confidence interval by the 2.5% and the 97.5% of this distribution of 5000 mean values.

Finally, the better registration strategies were determined on the basis of accepting results which had overlapping 95% confidence intervals with the lowest TRE sample quantities.

6.2.7 Validation Strategy

The errors computed for the better registration strategies on the training data will generally underestimate the expected errors on unseen data. Thus the expected registration error for similar breast deformations was determined by applying the better registration strategies to misaligned images of the five patients from the test set.

6.2.8 Sensitivity and Significance of Results

The sensitivity of the results to the partitioning of the data into training set and test set was examined by exchanging these sets, and by patient-wise leave-one-out tests for all 10 patients.

The sensitivity of the registration results to 7 factors was assessed. The factors were the patient, the image enhancement, the material model, the deformation scenario, the deformation magnitude, the registration type and the volume preservation regularization weight. The analysis was based on the model-free sensitivity index $S_i$ as described in Section 5.2.5 on page 161.

The effect that masking out locations outside the FEM mesh had on the registration was assessed. FEM deformed post-contrast images with background were created by employing a scattered data interpolation technique [Lee et al., 1997] based on a multilevel B-spline hierarchy for the background. The intensities within the FEM mesh were kept as before.

Statistical significance was tested against the probability that any of the 74 registrations could have produced TRE sample quantities that low by chance [Simon, 1997]. For this, the TRE sample quantities were averaged for the 12 FEM simulations for a given patient and registration configuration. The 74 average TRE sample quantities of each patient were
then randomly assigned to the 74 registration configurations by permutation. The 10 results randomly assigned to each registration configuration were then averaged and the minimal result for all registrations obtained in this way was recorded. A probability distribution of these minimal TRE sample quantities accomplished by chance was then generated by repeating this process 10000 times. The registration results were finally compared against this test statistics.

6.2.9 Analysis of Registration Error

Valuable insight about causes of registration errors can be gained by inspecting the spatial distribution of the target registration error. Furthermore, the influence of the image contrast change on the registration accuracy was studied. This analysis was based on comparing the results of registering three types of source images to the gold standard deformed post-contrast image \( (I_F) \). The source images were firstly the original pre-contrast image \( (I_1) \), secondly a simulated pre-contrast image \( (I_{1b}) \) generated by linearly blending the intensities of the pre-contrast lesion into the post-contrast image before adding noise and thirdly the post-contrast image with added noise \( (I_{2n}) \). Example slices of the difference between the post-contrast image \( I_2 \) and these source images are shown in Figure 6.8a on page 195.

6.3 Results

The registration performance of the registration strategies listed in Table 7.1 was investigated in this study. The assessment was based on calculating TRE at each voxel with respect to the simulated gold standard. The TRE distributions were summarized by the four sample quantities mean(mean\(TRE_B\)), mean(95\%TRE_B), mean(mean\(TRE_L\)) and mean(95\%TRE_L) as described in Section 6.2.5.

6.3.1 Selecting the Better Registration Strategies

Registrations were conducted for the datasets from 5 patients with 12 FEM simulations for each patient. These 60 pairs of deformed images are called the training set.

Firstly, the influence of the image similarity measure on the registration results was assessed for the training set, see Figure 6.3. Severe registration errors \( (\text{mean}(\text{mean}_{TRE_B}) > 65\text{mm}) \) occurred with sum of squared differences (SSD) and joint entropy (JE) for FFD registrations with 40mm and 60mm control point spacing, respectively. Two outliers (with a mean(mean\(TRE_B\)) > 1753mm) caused these high errors for SSD, while 29 cases (with a mean(mean\(TRE_B\)) > 262mm) contributed to the high errors for JE. The remaining cases achieved a mean(mean\(TRE_B\)) of 0.87mm (0.75mm) for SSD (JE). The misalignment of the
enhancing lesion increased substantially with sum of squared differences, cross correlation and correlation ratio for 10mm and 5mm FFD registrations. Smallest TRE sample quantities

![Graph showing TRE results for whole breast tissue and enhancing lesion with different registration methods.

Figure 6.3: Influence of the image similarity measure on the target registration error (TRE). The registrations were only optimized with respect to the image similarity (i.e. \( \mu = 0 \), see Table 7.1). The horizontal lines labelled 'Not reg.' show the initial errors. The seven bars for each registration configuration depict from left to right the results for sum of squared difference, cross correlation, correlation ratio \((A|B^T)\), correlation ratio \((B^T|A)\), joint entropy, mutual information and normalized mutual information. Top: average of mean TREs for all cases from the training set. Bottom: average of the 95th percentile TREs for all cases from the training set. Results were obtained for the whole breast tissue (left) and for the region of the enhancing lesion (right).
6.3 Results

were always achieved with either mutual information (MI) or normalized mutual information (NMI). Results were generally better for NMI than for MI (mean improvement 0.03mm, range [-0.02,0.22]mm). NMI was therefore used for every further registration.

Secondly, the four TRE sample quantities were calculated for all 74 registration configurations listed in Table 7.1 and for all cases of the training set. The results are depicted graphically in Figure 6.4. It can be seen that volume preservation was more important for FFD registrations with finer control point spacing and for the region of the enhancing lesion. Eleven better registration configurations were obtained when accepting results which had overlapping 95% confidence intervals with the lowest TRE sample quantities. These were 40mm single-resolution FFD registration with weight value $\mu \in [0.0,0.8]$ and 20mm single-resolution FFD registration with $\mu \in [0.7,0.8]$. These configurations reduced the mean

![Figure 6.4: Performance of the registration configurations described in Table 7.1 for the training set. The weight value $\mu$ balances the two objectives of the registration cost function defined in equation (6.1).](image-url)
registration error for the training set from 1.52mm to 0.48mm for the whole breast tissue and from 1.94mm to 0.35mm for the enhancing lesion. In comparison, affine registration reduced these values to 0.56mm and 0.44mm, respectively.

Volume-preserving FFD registrations with finer control point spacings (20mm single-resolution with $\mu \in [0.7,0.8]$ or 10mm multi-resolution with $\mu = 0.8$) produced generally better results for the whole breast tissue, for large deformations or for severe deformations (one-sided contact, two-sided contact, tension). The enhancing lesions, small deformations or smooth global deformations (point puncture, relaxation, regional displacement) were mostly best aligned with a 40mm control point spacing ($\mu \in [0,0.8]$).

Figure 6.5b shows examples of motion artifacts introduced by the FEM simulated deformations. These artifacts were greatly reduced after employing the registration configuration with the lowest mean(95\%TRE$_B$) for the larger deformations (Figure 6.5c). Some local registration failures at highly deformed regions can be observed for the configuration achieving the lowest registration errors for the whole breast tissue (Figure 6.5d). These are further increased by the configuration providing the lowest mean(95\%TRE$_L$) (Figure 6.5e). Yet, it is harder, if not impossible, to detect visually the more accurate alignment of the enhancing lesion in Figure 6.5e, since the true lesion enhancement pattern is unknown.

### 6.3.2 Validation on Test Data

Figures 6.6(c-e) show example slices after registering the test set for the same registration configurations as in Figure 6.5. Applying the 11 better registration configurations to the test set reduced the mean registration error from 1.40mm to 0.45mm for the whole breast tissue, and from 1.20mm to 0.32mm for the enhancing lesion. In comparison, affine registration reduced these values to 0.55mm and 0.44mm, respectively. The average 95th percentile registration error improved at the same time from 3.00mm to 1.10mm for the whole breast tissue and from 1.63mm to 0.46mm for the enhancing lesion. Affine registration reduced these values to 1.27mm and 0.57mm, respectively.

### 6.3.3 Sensitivity and Significance of Results

The same 11 better registration configurations were selected when exchanging the training and test set. These were also the only configurations which qualified as better configurations during patient-wise leave-one-out tests (20mm single-resolution FFD registration with $\mu = 0.7$ was selected 4 times and all other configurations were selected 10 times).

The sensitivity of the TRE sample quantities was assessed with respect to 5 patients from the training set, 3 source images ($I_1, I_{1b}, I_{2n}$), 2 material models (MM2, MM7), 6 deformation
Figure 6.5: Example slices from the training set showing difference (a) of pre- and post-contrast images; (b) after gold standard deformation; (c) after 10mm multi-resolution FFD registration with a volume preservation weight $\mu = 0.8$; (d) after 20mm single-resolution FFD registration with $\mu = 0.8$; (e) after 40mm single-resolution FFD registration with $\mu = 0.7$. The applied deformations were large regional displacement, common one-sided contact, large two-sided contact, common relaxation or large tension (left to right).

scenarios (see Section 6.2.3), 2 deformation magnitudes (common, large), 6 registration types (rigid, affine, s60, s40, s20, s10, m10, m5, see Table 7.1) and 2 volume preserving weight values ($\mu \in \{0, 0.8\}$). The TRE sample quantities for the whole breast were most sensitive to the deformation scenario followed by the deformation magnitude or the registration type, see Table 6.2. Registration errors for the enhancing lesions were mostly influenced by the type of source image and the patient, while the material model had the least influence.
Figure 6.6: Example slices from the test set showing difference (a) of pre- and post-contrast images; (b) after gold standard deformation; (c) after 10mm multi-resolution FFD registration with a volume preservation weight $\mu = 0.8$; (d) after 20mm single-resolution FFD registration with $\mu = 0.8$; (e) after 40mm single-resolution FFD registration with $\mu = 0.7$. The applied deformations were common regional displacement, large one-sided contact, common point puncture, large relaxation or common tension (left to right).

Registrations which included the image content outside the FEM mesh were best aligned when decreasing the weights of the better registration strategies by 0.1, i.e. 40mm single-resolution FFD registration with $\mu \in [0,0.7]$ and 20mm single-resolution FFD registration with $\mu \in [0.6,0.7]$.

Only 20mm single-resolution FFD registration with $\mu=0.8$ achieved statistically significantly better results than chance at the 5% level for all TRE sample quantities. The other
6.3 Results

<table>
<thead>
<tr>
<th>Factors</th>
<th>$S_i$ in % for</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean $T_R E_{B}$</td>
</tr>
<tr>
<td>Patient</td>
<td>1</td>
</tr>
<tr>
<td>Source image</td>
<td>5</td>
</tr>
<tr>
<td>Material model</td>
<td>0</td>
</tr>
<tr>
<td>Deformation scenario</td>
<td>18</td>
</tr>
<tr>
<td>Deformation magnitude</td>
<td>9</td>
</tr>
<tr>
<td>Registration type</td>
<td>13</td>
</tr>
<tr>
<td>Weight value</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 6.2: Sensitivity analysis of TRE sample quantities. The sensitivity index $S_i$ measures the average variance reduction of the TRE sample quantities when fixing the listed factor. Influential factors are indicated by large positive $S_i$ values.

better registration strategies either failed to provide statistically significant improvements for $\text{mean}(95\% T_R E_{B})$ (40mm single-resolution FFD with $\mu \in [0, 0.8]$) or for the enhancing lesion (20mm single-resolution FFD with $\mu = 0.7$).

6.3.4 Analysis of Registration Error

The influence of the image enhancement on TRE was studied by registering three types of source images ($I_1$: original pre-contrast image, $I_{1b}$: simulated pre-contrast image, $I_{2n}$: post-contrast image with added noise) to the FEM deformed post-contrast image ($I_F$). The average registration errors for these image pairs are depicted in Figure 6.7. Images which differed initially only by noise ($I_{2n}$) were best aligned with FFD registrations of 20mm or finer control point spacing. The introduction of image intensity changes due to contrast agent (within the whole breast region for $I_1$ and within the region of the enhancing lesion for $I_{1b}$) caused a substantial error increase for these registrations. Volume preservation ($\mu = 0.8$) lessened this tendency. More flexible registration types (more degrees of freedom, $\mu = 0$) were more effective in reducing maximum registration errors for the whole breast tissue for images which differed outside the enhancing lesion only by noise ($I_{2n}$ and $I_{1b}$).

Figures 6.8(b,c) show the spatial distribution of the average TRE for the three types of source images. In this case, the average TRE denotes the mean of the registration results for the 12 FEM simulations for a patient. Increased TRESs because of image intensity changes can be observed in Figure 6.8b for the most flexible registration configuration (5mm multi-
Figure 6.7: Influence of image enhancement on registration results. Three different source images ($I_1$: pre-contrast image, $I_{1b}$: post-contrast image with blended pre-contrast lesion, $I_{2n}$: post-contrast image with added noise) were registered to the FEM deformed post-contrast image ($I_F$) for each registration strategy (labels on x-axis, see Table 7.1). Registration results are summarized by the average of the mean target registration errors (TREs) (bars), the average of the 95th percentile TREs (lines) and the average of the maximum TREs (symbols) for all cases from the training set. Results were evaluated over the whole breast tissue (top) and within the enhancing lesion (bottom).
Figure 6.8: Example slices for three types of source images of patient 5, namely pre-contrast image ($I_1$), post-contrast image with blended pre-contrast lesion ($I_{1b}$), post-contrast image with added noise ($I_{2n}$). (a) Difference of post-contrast image ($I_2$) and source image; (b, c) spatial distribution of average target registration error (TRE) for all 12 registrations for this patient (black: TRE=0mm, white: TRE ≥ 1.5mm) after (b) 5mm multi-resolution FFD registration without constraining for volume preservation ($\mu = 0$, $m50$). (c) 20mm single-resolution FFD registration with $\mu = 0.8$ ($s20\mu0.8$); (d, e) average of local volume changes (dV) of all 12 registrations for this patient from dV ≤ -10% (black) to dV ≥ 10% (white) after (d) $m5\mu0$ and (e) $s20\mu0.8$. 
resolution FFDs with $\mu = 0$). Problems with adapting to local sharp deformations at the skin surface are visible for the configuration achieving the lowest TREs for the whole breast tissue, see Figure 6.8c. A tendency of the more flexible registrations to shrink the region of the enhancing lesion is illustrated in Figure 6.8d. Coarser control point spacing and penalizing volume changes prohibited this tendency (Figure 6.8e).

### 6.3.5 Computational Costs

Best registrations were obtained most rapidly with 40mm single-resolution FFD configurations where results were obtained on average within 26 minutes on a 1.8 GHz Athlon processor with 1GByte 1.33 MHz SD RAM memory. 20mm single-resolution FFD registration required on average 1 hour 21 minutes.

### 6.4 Summary

This chapter described an extensive validation study of a non-rigid registration method for the intra-visit registration of contrast-enhanced MR breast images. The assessment was based on simulating biomechanically plausible breast deformations as a gold standard. Firstly, image pairs of ten patients were selected where no motion between pre- and post-contrast image was visible. Secondly, twelve gold standard deformations were generated for each patient employing a configuration of the biomechanical breast model which was determined as suitable in Chapter 5. Thirdly, the better registration protocols were identified using a training set which included all simulated deformations of five of the patients. Finally, the registration accuracy of these better registration strategies for similar deformations was established on a test set consisting of deformations for the other five patients. Furthermore, the sensitivity of the results to various factors was assessed and the results were tested for statistical significance. Lastly the influence of the image enhancement on the registration errors was studied.

Two components mainly contributed to the registration errors, namely image intensity changes due to the contrast agent and severe deformations. Registrations which permitted very flexible transformations (i.e. with a small control point spacing and a weak volume preservation constraint) compensated well for severe deformations, yet errors increased in regions influenced by changes in image intensity. Restricting the registration by increasing the control point spacing or by promoting volume preservation provided better alignment of the enhancing lesion but worsened adaption to severe deformations. The better registration strategies for both regions of interest were 40mm single-resolution FFD registration with $\mu \in [0,0.8]$ and 20mm single-resolution with $\mu \in [0.7,0.8]$. These strategies achieved on average
expected mean (95th percentile) errors of 0.45mm (1.10mm) for the whole breast tissue and of 0.32mm (0.46mm) for the enhancing lesion. The best results were obtained most rapidly in 26 minutes with the 40mm registrations, while the 20mm configuration was better suited for larger and more severe deformations.

Volume preservation was more important for registrations with finer control point spacing. The optimal volume preserving weight values were between 0.7 and 0.9 for most FFD registration configurations. Registrations which included the image content outside the FEM mesh were best aligned when decreasing these optimal weights by 0.1. The exact optimal weight value for volume preservation is unlikely to be directly applicable to other registration configurations or implementations. Anatomically, volume is (within a certain tolerance) locally conserved within the breast during MR mammography. In contrast, the employed regularization penalized average local volume change. Thus, registration may improve if volume can be preserved locally without causing early termination (locking).

The FEM and FFD transformations are related in the sense that both are defined by a finite number of displacement vectors. They differ substantially, however, due to the number and position of these displacement vectors and due to their interpolation functions\(^1\). Even a perfect registration, where the FFD parameters are optimized to minimize the TRE with regard to a known FEM transformation, will therefore have a residual error. This optimal residual error will be lower for transformations with more degrees of freedom. Such an error decrease occurred for maximum registration errors of the whole breast tissue for images which differed outside the enhancing lesion only by noise. Otherwise, however, the contrary happened, although the FEM transformations had on average over 7 times more degrees of freedom than the most flexible registration tested in this study\(^2\). This indicates that limited image content and change in image contrast rather than bias or simplicity of the biomechanical breast models have caused this effect.

The role of the biomechanical breast models for the validation was to simulate breast deformations which are likely to occur during DCE MR mammography rather than predict accurately the deformation of a specific patient. This requires a plausible elastic breast model and plausible deformation scenarios which include the range of local and global deformations

\(^1\)Piecewise quadratic shape interpolation of irregularly sized 10-noded tetrahedral FEM elements versus \(C^2\) continuous B-spline shape approximation of equally sized 8-noded brick elements.

\(^2\)FEM meshes consisted on average of 95133 mesh nodes. Breast regions were on average under the influence of 13095 control points for 5mm multi-resolution FFD registrations. Each mesh node and control point had 3 degrees of freedom.
which will be encountered in practice. Biomechanical models were able to predict the position of internal landmarks for 20% in-vivo breast compressions with an average accuracy of 2.1mm, see Chapter 5. Furthermore, registration results were least sensitive to the choice of material model. The biomechanical models used a different interpolation function than the registration to avoid introducing any bias. Sliding breast tissue components were, however, not modelled. Increased registration errors can be expected at such sliding interfaces due to the continuous nature of the transformation models used in the registrations. The simulated breast deformations included localized, smooth and severe deformations. The most unlikely of these deformations during DCE MR mammography were point puncture and two-sided contact. Yet, the better registration strategies yielded similar results whether these scenarios were excluded or not. The simulated displacement magnitudes reflect that the aim was to assess the registration performance for breast deformations and not for major global patient repositioning.

In this study, the pre-contrast image was registered to the deformed second post-contrast image. The use of other image pairs from the dynamic series may result in different optimal registration strategies due to altered enhancement patterns. Similarly, the study’s conclusion may not apply to other types of registration, e.g. other organs or without contrast change.

The validation was based on five patients for the training set and five patients for the test set, due to the relatively small number of image pairs without detectable motion. Exchanging the training and the test set produced the same better registration configurations and similar validation results. These were also the only configurations qualifying for the better configurations during patient-wise leave-one-out tests. This robustness is in contrast to the strong influence individual patients had on the registration error for the enhancing lesion. Hence one can be confident that the range of breast compositions across the population was captured and that the main results are generally applicable.

Only one of the better registration strategies provided results which were statistically significantly better than chance. Significance of other configurations was most likely not detected because of the small number of patients. The achieved significance indicated that the observed registration improvement was not an artifact.

The availability of a gold standard not only provided a means to optimize and validate the non-rigid registration, but also to gain insight into the registration problem, which will help in developing better registration approaches for contrast-enhanced images in general. For example, this study indicated that the difference in image intensity due to contrast agent caused the accuracy to degrade as the control point spacing was reduced. This problem
was mainly caused by the ability of flexible registrations to deform the image locally and by the improvement of the image similarity measure when altering the enhanced lesion. Such an effect is likely to hold in general when registering contrast-enhanced image pairs with algorithms of similar flexibility and optimization objectives. Note that this problem cannot be detected by visual inspection alone, since the true pattern of contrast change is unknown.

A subset of the validation dataset has been used to assess the accuracy of a fluid based registration method for unenhanced MR image pairs ($I_{2n}$) [Crum et al., 2005]. This fluid based registration method performed slightly worse than a 20mm single-resolution FFD registration, which was not constrained to preserve volume since the images were not contrast-enhanced. Registration methods from other groups have not yet been assessed with this validation dataset, but are welcomed.
Chapter 7

Computer Aided Diagnosis System for DCE MR Mammography

This chapter presents a computer aided diagnosis (CAD) system for DCE MR mammography and describes its classification performance. The CAD system is based on the best registration method found in Chapter 6, a semi-automatic segmentation method devised and tested in this chapter, the extraction of promising features as described in Chapter 4 and logistic regression classifiers as discussed in Chapter 4. The content of this chapter is based on and extends [Tanner et al., 2004].

7.1 Introduction

A framework for a CAD system for DCE MR mammography has been proposed in Chapter 4. Its main components are image registration, segmentation refinement, feature extraction and classification as shown in Figure 4.1 on page 124.

An image registration method for DCE MR mammography was evaluated in Chapter 6 and the best registration configuration, with respect to the target registration error, was identified. Its importance as a component of a CAD system will be assessed in this chapter.

Manual segmentation of DCE MR mammograms is time consuming and complex since the whole image sequence needs to be considered. Only a few segmentation methods have been proposed for DCE MR mammography and their approaches are limited. Instead a constrained maximum a posteriori strategy is devised in Section 7.2.3 to extract the most probable lesion for a user-provided coarse manual outline. The segmentation results were visually assessed and the best configuration was employed for the CAD system.

Many features could be extracted from the lesion segmentations. Yet, the number of observations, i.e. the number of lesions assessed, restricts the number of feature candidates which should be evaluated when building a classifier [Hair et al., 1998; Schulerud and Albregtsen, 2004]. A three stage approach for feature selection was therefore employed. Promising features were firstly identified by reviewing the literature, see Section 4.3. Thereafter, only the least correlated features were selected as candidates for logistic regression, such that the recommended ratio between the number of observations and features was achieved. Then the final selection was done by the classifier.
Linear discriminant and logistic regression analysis are two of the most widely used classifiers. These were reviewed in Section 4.4 and logistic regression analysis was identified as the favoured model, mainly because of its greater robustness to violations of data distribution assumptions.

Classifiers were built using features obtained from a manual segmentation or from a refined segmentation of the original images or the images after registration. These were compared to classifiers based on the cumulative scores from radiologists. The classification performance was quantified by leave-one-out receiver operating characteristics (ROC) analysis.

The remainder of this chapter is organized as follows. Section 7.2 describes the datasets, the registration configurations, the segmentation refinement method, the features and the classification approach. The results are presented in Section 7.3 and the main findings are discussed in the final Section 7.4.

7.2 Materials and Methods

7.2.1 Datasets

Forty-four patients were randomly selected from the symptomatic database of the UK multicentre study of MRI screening in women at genetic risk of breast cancer (MARIBS). The patients had in total 18 benign and 38 malignant histologically proven lesions. The images came from seven centres of the MARIBS study and were all acquired according to an agreed protocol (3D gradient echo sequence on a 1.5T MR system with TR=12ms, TE=5ms, flip angle=35°, field of view of 340mm, 1.33×1.33×2.5mm voxel size, coronal slice orientation, 90s acquisition time, 0.2mmol Gd-DTPA, see [Brown et al., 2000]).

The lesions were manually segmented by the author according to the annotation from an experienced radiologist. The radiologist had access to the clinical reports to ensure that the correct lesion was annotated. The author segmented the lesions by employing region growing techniques from ANALYZE [AnalyzeDirect Inc., 2005] with manual corrections where necessary. Generally, the same intensity threshold was applied to all slices while the seed voxel was moved. Crude manual outlines were simulated by approximating the manual segmentation by an ellipse on each 2D slice.

7.2.2 Image Registration

In order to assess the influence of image registration on the performance of the CAD system, the image sequence was registered in four different ways as summarized in Table 7.1. The non-rigid registration was based on the multi-resolution free-form deformation (FFD) approach employing B-spline subdivision as described in [Rueckert et al., 1999b] and reviewed
in Chapter 2, see page 80. The choice of registration methods was guided by the results from Chapter 6, where a 20mm volume-preserving non-rigid registration (NR20V) was amongst the best configurations. The other strategies were no registration, rigid registration and 10mm non-rigid registration as used previously [Rueckert et al., 1999b; Denton et al., 1999; Sonoda, 2003].

7.2.3 Segmentation Refinement

This section describes the devised segmentation refinement method and its evaluation.

7.2.3.1 Introduction

The segmentation refinement aimed to extract the most probable 3D lesion object of the DCE MR image sequence from the data provided by a manual outline and prior knowledge about the segmentation process. The problem was posed as a two-class classification problem where the training data were provided by manual outline. The segmentation decision was based on a constrained maximum a posteriori probability (MAP) estimation in order to account for the imbalanced number of lesion and background voxels. The class conditional probability density functions were directly estimated from the temporal domain of the data samples. Sparsely sampled distributions were avoided by reducing the temporal dimensions with principle component analysis.

Manually segmented lesions generally include regions of non-enhancing tissue (like small heterogeneities, necrotic centres or fatty regions). A MAP estimation solely based on the temporal domain would therefore lead to misclassifications. Instead, the MAP estimation was rearranged such that the ratio of the a priori class probabilities can be viewed as a threshold for the likelihood ratio. The segmentation process was then modelled by extracting the

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unreg.</td>
<td>No registration, i.e. unregistered image sequence</td>
</tr>
<tr>
<td>Rigid</td>
<td>Rigid registration</td>
</tr>
<tr>
<td>NR20V</td>
<td>Single-resolution FFD registration after rigid registration with a control point spacing of 20mm and a weight value of $\mu = 0.7$ in equation (6.1) to penalize volume changes</td>
</tr>
<tr>
<td>NR10</td>
<td>Single-resolution FFD registration after rigid registration with a control point spacing of 10mm and without penalizing volume changes ($\mu = 0$)</td>
</tr>
</tbody>
</table>

Table 7.1: List of assessed registration strategies. normalized mutual information was used as image similarity measure in all cases.
biggest connected and filled 3D lesion object for a given thresholded likelihood ratio map. The lesion candidate with the highest average a posteriori probability, which changed in size by less than a given limit, was then selected as the most probable lesion. No assumptions were made about the edge strength, the shape or the enhancement profile of the lesion in order to avoid removing valuable information for the discrimination of benign and malignant lesions.

**7.2.3.2 Data-Preprocessing**

The image data was preprocessed for segmentation refinement by subtracting the mean of the two pre-contrast images from each post-contrast image. For each lesion, the 3D region of interest (ROI) was defined as the rectangular box which contains the 3D manual outline plus a margin of at least 7mm in each direction. This ROI size was chosen in order to speed-up the processing, to have more homogeneous properties of the background voxels and to ensure that manual segmentation errors are contained within the ROI.

Many of the multispectral segmentation algorithms assume that the intensity distributions of the separate objects can be approximated by multivariate Gaussian distributions [Pham et al., 2000]. There is, however, no reason to expect that the temporal data of DCE MR mammograms conform to this assumption. Density estimations were therefore performed with Gaussian kernels and a bandwidth selected according to [Silverman, 1986]. The dimensionality of the preprocessed data was reduced through principle component analysis to reduce sparseness.

**7.2.3.3 Constrained Maximum a Posteriori Probability Approach**

The segmentation refinement aimed to extract the most probable connected lesion object of a 3D ROI for a given manual outline. Assuming equally likely image features $x$ and taking the a priori class probability $P(c_k)$ for class $c_k$ into account, yields the most probable segmentation refinement by maximizing the a posteriori probabilities, i.e. $\max_k P(c_k | x) = P(x | c_k) P(c_k)$, where $P(x | c_k)$ was estimated from manual outline. For a two class problem the discrimination function $y(x)$ can be written as

$$y(x) = \frac{P(x | c_1)}{P(x | c_2)} \quad \text{with} \quad x \in \left\{ \begin{array}{ll} c_1 & \text{if } y(x) > \theta \\ c_2 & \text{otherwise} \end{array} \right. \quad \text{where} \quad \theta = \frac{P(c_2)}{P(c_1)}.$$  \hspace{1cm} (7.1)

Equation (7.1) emphasizes that the ratio of the a priori probabilities act as a threshold ($\theta$) on the likelihood ratio. Instead of estimating $\theta$ from the number of lesion and background voxels in the manual outline, $\theta$ was used for implicitly incorporating prior knowledge about the segmentation process as described next.
Firstly, the biggest connected object was extracted for a given threshold $\theta$. This assumed that one connected lesion was manually outlined per ROI. Morphological closing and hole filling operations were then applied to model the observation that manually segmented lesions generally include non-enhancing regions. Thereafter, a set of lesion candidates was generated by varying the threshold $\theta$. Assuming that the manual outline is similar in size to the actual lesion object, all candidates which had a volume change of less than a certain percentage compared to the manual outline were then selected. From this subset, the object with the maximum average a posteriori probability for the whole lesion was finally chosen.

Note that this refinement method can easily be changed into an interactive segmentation tool, where the user is given control over the threshold $\theta$. In this work, however, refinements were required for unregistered images as well as images registered differently. Segmentation refinement was therefore retrospectively employed for segmentation propagation to reduce the radiologist’s work load. The segmentation performance of the configurations listed in Table 7.2 was compared.

### 7.2.3.4 Segmentation Assessment

Segmentation methods are commonly evaluated by comparing them with manual segmentations by means of the overlap measure $O$,

$$O = \frac{V(A \cap B)}{V(A \cup B)}, \tag{7.2}$$

where $A$ and $B$ are two segmented regions; $A \cap B$ ($A \cup B$) are the intersection (union) of region $A$ and $B$; and $V(C)$ is the volume of region $C$. This evaluation method requires manual segmentations for all cases, which can be very time consuming when the influence of several registration methods need to be assessed. Furthermore, the overlap values do not

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Man.</td>
<td>Manual segmentation based on the subtraction of an unregistered image pair</td>
</tr>
<tr>
<td>ML</td>
<td>Maximum likelihood decision, i.e. $\theta = 1$ in (7.1)</td>
</tr>
<tr>
<td>MAP</td>
<td>Maximum a posteriori probability decision, i.e. $\theta = V(c_2)/V(c_1)$ in (7.1), where volume $V(c_k)$ was estimated from the manual outline</td>
</tr>
<tr>
<td>$C_p$</td>
<td>Constrained maximum a posteriori probability decision, extracting the connected filled lesion which changed volume by less than $p%$ while maximizing the average a posterior probability, i.e. $\theta = fV(c_2)/V(c_1)$ in (7.1) where $f \in [-p,p]/100$ and $\max_f\text{mean}(P(c_2</td>
</tr>
</tbody>
</table>

Table 7.2: List of assessed segmentation strategies.
necessarily reveal the quality of the segmentation.

Instead a two stage assessment was used. Firstly, the refined and manual segmentations were compared for the original images based on the overlap measure. Secondly, the author blindly ranked the refined segmentations according to their visual appearance for the 4 registration methods as described in Table 7.1 and for a subset of 4 refinement strategies. This was achieved, by automatically presenting the author with figures showing the segmentations overlayed on the images, with segmentation strategies randomly assigned to rows 2 to 5, as illustrated in Figure 7.5. Rows 1 and 6 always showed the initial outline and the manual segmentation, respectively. Figures were generated for all slices of the lesion ROI.

The ultimate aim of the segmentation refinement is to provide a lesion outline which improves classification. While one could optimize the whole process of registration, segmentation and feature extraction to find a configuration which provides the best classification, such an approach would require an enormous dataset to have enough statistical power for multiple comparisons. Instead, the segmentation strategy which had the best visual appearance according to the ranking was employed for all classifications.

7.2.3.5 Summary

This section described the segmentation refinement method and its evaluation. Having established a lesion segmentation, features can be generated from it as described next.

7.2.4 Features

Twenty-five promising features for MR breast lesion classification were identified in Section 4.3, see Table 4.6 on page 142. These were extracted from the segmented lesions.

The feature set was further reduced to the 11 least correlated features of the manual segmentation using hierarchical clustering [Theodoridis and Koutroumbas, 1999]. This was required, since a minimum of 5 observations is recommended for each feature candidate during discriminant analysis [Hair et al., 1998].

Uniformly distributed features were randomly generated to assess the classification performance solely achieved by the feature selection process and leave-one-out test procedure.

The lesion appearance on the original images was scored by two radiologists according to the MARIBS protocol [Brown et al., 2000], see Table 7.3. Generally, the images were first rated at the local study centre and then at another study centre. This meant that the same two radiologist could not be used throughout the study. The radiologists were not necessarily blind to the clinical symptoms. The classification performance of the cumulative scores from the two centres, named Score1 and Score2, was assessed in this work.
### Feature

<table>
<thead>
<tr>
<th>Feature</th>
<th>Mutual exclusive feature description and scored points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphology</td>
<td>Lobulated, Poorly well-defined, Spiculated or focal branching</td>
</tr>
<tr>
<td>Points</td>
<td>0, 5, 10</td>
</tr>
<tr>
<td>Pattern of enhancement</td>
<td>Centrifugal/no enhancement, Heterogeneous Ring-like enhancement OR minimal enhancement/homogeneous enhancement</td>
</tr>
<tr>
<td>Points</td>
<td>0, 3, 5</td>
</tr>
<tr>
<td>$K_{rme}$</td>
<td>$&lt;40%$, 40-60%, $&gt;60%$</td>
</tr>
<tr>
<td>Points</td>
<td>0, 2, 4</td>
</tr>
<tr>
<td>$K_{mitr}$</td>
<td>$&lt;25%/min$, 25-49%/min, $&gt;50%/min$</td>
</tr>
<tr>
<td>Pattern of contrast</td>
<td>Monotonic Peak signal Decrease signal increase by 3min is directly maintained after peak</td>
</tr>
<tr>
<td>Points</td>
<td>0, 4, 8</td>
</tr>
</tbody>
</table>

Table 7.3: MARIBS scoring system for 3D image sequences. The cumulative score is determined by adding the points scored by the 5 features and hence can take values between 0 and 35. $K_{rme}$ denotes the relative maximum enhancement calculated as $\bar{I}_{max}/\bar{I}(0) - 1$, where $\bar{I}(k)$ denotes the mean intensity of the lesion for image $k$, $\bar{I}_{max} = \max_k \bar{I}(k)$ and image 0 is the pre-contrast image. $K_{mitr}$ represents the maximum intensity time rate defined as $(\bar{I}_{max} - \bar{I}(0))/t_{max}$ where $t_{max}$ is the time between contrast agent administration and when $\bar{I}_{max}$ occurred.

### 7.2.5 Classification

Logistic regression analysis, as described in Section 4.4.3 on page 145, was employed for all classifiers. Multivariate models were built using a combination of forward selection and backward elimination with a probability threshold of 0.20 and 0.25, for the likelihood ratio chi-square test as recommended in [Hosmer and Lemeshow, 2000].

The classification performance was assessed by two measures. Firstly, the t-test probability, which denotes the probability that the mean value for benign and malignant lesions is the same as determined by independent Student t-tests. The second measure is the area under the ROC curve (AUC) from leave-one-out test as defined in Section 4.4.4.2, page 147. Leave-one-out AUC values were tested for statistically significant differences using the ROCKIT program from [Metz et al., 1998].
7.3 Results

This section summarizes the results from the segmentation, the feature selection and the classification.

7.3.1 Segmentation

Figures 7.1 to 7.4 illustrate for an example slice the main processing steps of the 3D segmentation refinement method. The sequence of 4 subtraction images in Figure 7.1 was reduced to two images by employing principle component analysis as shown in Figure 7.2. The first and second eigenvector explained 73% and 23% of the variation in the data for this example. The class conditional probability distribution functions were then estimated from the crude outline of the principle component images, see Figure 7.3. Figure 7.4 depicts the segmentation result when using a maximum a posteriori criterion, i.e. when thresholding the likelihood ratio of equation (7.1) by the ratio of the a priori probabilities.

The background (lesion) distributions of the original data were statistically significantly different at 5% level based on the Lilliefors test from Gaussian normal distributions in 100% (74%) of all cases. The first two principle components of the preprocessed enhancement curves describe on average 97% of the variation in the data (range [83,100]%). The background (lesion) distributions of these two components were statistically significantly different from Gaussian normal distributions in 100% (65%) of all cases.

Figure 7.5 shows for an example slice the refined segmentations of an unregistered image sequence. It can be observed that the ML refinement (contour in row 5) of the initial outline (row 1) overestimated the extent of the lesion when compared with the manual segmentation (row 6). The MAP (row 3) and C0 (row 4) refinements underestimate the lesion slightly. The C50 (row 2) refinement was closest to the manual segmentation for this slice. The overlap, as defined in equation (7.2), between the 3D manual segmentation and the 3D refinement was 0.60, 0.49, 0.57 and 0.32 for C50, MAP, C0 and ML, respectively.

Figure 7.6 depicts the segmentation refinement results for the same slice after 20mm volume-preserving non-rigid registration. The pre-contrast image has changed, the enhancing lesion shrunk and the manual segmentation from the original images no longer fits the lesion. Using the same initial outline as for the original images, the C50, MAP and C0 refinements (rows 2-4) produced acceptable segmentations, while the ML refinement (row 5) outsized the lesion. The overlap between the 3D manual segmentation and the 3D refinement reduced to 0.31, 0.32, 0.37 and 0.17 for C50, MAP, C0 and ML, respectively. This case had the smallest overlap of all 56 lesions which were refined using the C50, C0 or ML criterion after 20mm
Figure 7.1: Example slice showing (left to right) difference images after subtracting the pre-contrast image from the 1st, 2nd, 3rd and 4th post-contrast image. The overlayed contour shows the crude outline.

Figure 7.2: Result of applying principle component analysis to the 3D image used in Figure 7.1. The left graph shows the first two eigenvectors. The example slice after the principle component transformation is shown on the right.

Figure 7.3: The class conditional probabilities $P(x_1, x_2 | c_k)$ from the principle component images and the crude outline in Figure 7.2 for (left) the background and (right) the lesion.

Figure 7.4: Example slice with contour showing the maximum a posteriori segmentation.

volume-preserving non-rigid registration. It is a striking example where the motion artifact has mimicked an enhancing lesion on the subtraction images.
Figure 7.5: Segmentation example for the original image sequence. The example slice showing (left to right) pre-contrast image, and difference images after subtracting the pre-contrast image from the 1st, 2nd, 3rd and 4th post-contrast image. Overlayed contours show (top to bottom) initial outline, refinement of initial outline by C50, MAP, C0, and ML criterion (see Table 7.2), and manual segmentation.

Figure 7.6: Segmentation example after 20mm volume-preserving non-rigid registration.
Figure 7.7 shows how the refinement criteria and the registration strategy affected the overlap measure. Applying a threshold to keep volume changes to a minimum (C0) provided on average the largest overlap for each registration strategy. Overlaps generally decreased with the more flexible registrations and segmentation schemes.

The manual segmentation was based on the original images and hence was an inadequate reference when motion artifacts were present, as for example in Figure 7.6. Manual segmentations would be required after all registrations. Instead, the visual appearance of 4 segmentation strategies, namely C0, C50, MAP and ML, was blindly ranked by the author. The results of assessing the segmentation quality are shown in Table 7.4. The maximum likelihood (ML) approach provided in most cases a worse segmentation. The quality of the other three segmentation methods was quite similar with C0 being slightly better. The C0 segmentation strategy was therefore used for the further tests.
<table>
<thead>
<tr>
<th>Method A</th>
<th>Method B</th>
<th>A better than B</th>
<th>A==B</th>
<th>A worse than B</th>
</tr>
</thead>
<tbody>
<tr>
<td>C0</td>
<td>ML</td>
<td>215</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>C50</td>
<td>ML</td>
<td>212</td>
<td>5</td>
<td>7</td>
</tr>
<tr>
<td>MAP</td>
<td>ML</td>
<td>214</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>C0</td>
<td>MAP</td>
<td>46</td>
<td>147</td>
<td>31</td>
</tr>
<tr>
<td>C50</td>
<td>MAP</td>
<td>42</td>
<td>141</td>
<td>41</td>
</tr>
<tr>
<td>C0</td>
<td>C50</td>
<td>40</td>
<td>156</td>
<td>28</td>
</tr>
</tbody>
</table>

Table 7.4: Results of blindly ranking the segmentation quality of four segmentation methods (C0, C50, MAP, ML, see Table 7.2). A total of 56 lesions registered in 4 different ways (Unreg., Rigid, NR20V, NR10, see Table 7.1) were assessed.

### 7.3.2 Feature Extraction

Twenty-five promising features, as defined in Table 4.6 on page 142, were extracted from each lesion. The 11 least correlated features were selected in order to have a minimum of 5 observations per feature. The features were first clustered as shown in Figure 7.8. It can be observed that features which were positively or negatively correlated by more than 0.41 (equal to a distance of less than 0.59) had to be considered for removal. For clusters with

![Figure 7.8](image-url)

Figure 7.8: Hierarchical clustering of 25 features (defined on page 142) derived from the manual segmentation. The distance between two features is given by $1 - |r|$, where $r$ is the correlation coefficient. In order to get compact clusters, a complete link algorithm was used, where the distance of a new cluster is given by the maximum distance of its subclusters. The dashed line represents the threshold value for 11 clusters.
more than two features, the feature which was most correlated with the features of the cluster was chosen as the cluster representative. The least normal distributed feature was removed from clusters consisting of two features. The selected features are summarized in Table 7.5.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Category</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_{comp}$</td>
<td>Shape</td>
<td>Compactness</td>
<td>132</td>
</tr>
<tr>
<td>$M_{vmmg}$</td>
<td>Margin</td>
<td>Variance of maximum margin gradient</td>
<td>134</td>
</tr>
<tr>
<td>$K_{slt}$</td>
<td>Kinetic</td>
<td>Slope factor</td>
<td>138</td>
</tr>
<tr>
<td>$K_{sde}$</td>
<td>Kinetic</td>
<td>Standard deviation of enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{rie}$</td>
<td>Kinetic</td>
<td>Relative initial enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$K_{mwo}$</td>
<td>Kinetic</td>
<td>Maximum washout</td>
<td>138</td>
</tr>
<tr>
<td>$K_{ser}$</td>
<td>Kinetic</td>
<td>Signal enhancing ratio</td>
<td>138</td>
</tr>
<tr>
<td>$K_{lpe}$</td>
<td>Kinetic</td>
<td>Late peripheral enhancement</td>
<td>138</td>
</tr>
<tr>
<td>$T_{sue}$</td>
<td>Texture</td>
<td>Sum average</td>
<td>139</td>
</tr>
<tr>
<td>$T_{imuv}$</td>
<td>Texture</td>
<td>Index of maximum uptake variance</td>
<td>139</td>
</tr>
<tr>
<td>$O_{volume}$</td>
<td>Other</td>
<td>Lesion size</td>
<td>140</td>
</tr>
</tbody>
</table>

Table 7.5: Overview of the 11 least correlated features.
### 7.3 Results

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Feature description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Score1</td>
<td>Cumulative score, as defined in Table 7.3, from local radiologist</td>
</tr>
<tr>
<td>Score2</td>
<td>Cumulative score, as defined in Table 7.3, from second radiologist</td>
</tr>
<tr>
<td>Man.</td>
<td>Features obtained from the manual segmentation of the original images</td>
</tr>
<tr>
<td>Unreg.</td>
<td>Features obtained from the refined segmentation of the original images</td>
</tr>
<tr>
<td>Rigid</td>
<td>Features obtained from the refined segmentation after rigid registration</td>
</tr>
<tr>
<td>NR20V</td>
<td>Features obtained from the refined segmentation after 20mm volume-preserving non-rigid registration as defined Table 7.1</td>
</tr>
<tr>
<td>NR10</td>
<td>Features obtained from the refined segmentation after 10mm non-rigid registration as defined Table 7.1</td>
</tr>
<tr>
<td>Random</td>
<td>Randomly generated features with uniform distribution called $R_i$</td>
</tr>
</tbody>
</table>

Table 7.6: Description of classifiers. Segmentations were refined using the volume constrained approach, i.e. C0 as defined in Table 7.2. Feature sets included either all promising 25 features (Table 4.6, page 142) or the subset of the 11 least correlated features (Table 7.5). Logistic regression was used to create the classifiers.

#### 7.3.3 Classification

This section describes the classification performance of single and combined features. Classifiers were built for the radiological scores and for 6 feature extraction methods, as described in Table 7.6, employing logistic regression models.

#### 7.3.4 Individual Feature Performance

The classification performance of the individual features is described in this section, to allow comparison with the performance after combining features. The evaluation was based on both the t-test probability, i.e. the probability that the mean feature value for benign and malignant lesions is the same, and on the area under the ROC curve (AUC) from leave-one-out tests. Results were summarized by the mean values for the best 5, 10, 15, 20 and 25 features to illustrate the general trend.

Figure 7.9 shows the effect that segmentation refinement and registration had on the average t-test probability. Here segmentation refinement shows a slight reduction of the t-test probabilities (Man. versus Unreg.) and lowest average t-test probabilities were generally achieved for features extracted from the refined segmentations of rigidly registered images (Rigid).

The class means of six kinetic features ($K_{gwo}$, $K_{ser}$, $K_{mwo}$, $K_{rmwo}$, $K_{tmx}$, $K_{rie}$) and one margin feature ($M_{emg}$) were statistically significantly different at the 5% level for benign
and malignant lesions for all methods, see Table 7.7. The t-test probabilities of 3 texture features ($T_{var}, T_{sue}, T_{ent}$) became significant after rigid or 20mm volume-preserving non-rigid registration (NR20V). The only other feature which had a significant t-test probability was $T_{imuv}$, when extracted from the refined segmentation of the original images (Unreg.). The set of the 11 least correlated features (marked with $\rightarrow$ in Table 7.7) appears to be unbiased to the classification performance.

Registration and segmentation refinement also improved the average leave-one-out area under the ROC curve (AUC) of the individual features, see Figure 7.10. Maximum leave-one-out AUC values were achieved by rigid registration or 20mm volume-preserving non-rigid (NR20V) registration.

The leave-one-out AUC value for the individual features are recorded in Table 7.8. As expected, the sequence is somewhat different from the sequence in Table 7.7. The signal enhancing ratio feature ($K_{ser}$) clearly provided the best leave-one-out AUC values ranging from 0.822 to 0.864. This was not reflected in the t-test probabilities, where the general washout feature ($K_{gwo}$) was on average slightly better than $K_{ser}$. The leave-one-out performance of the time to maximum enhancement feature ($K_{tmx}$) was notably worse, this was possibly due to its discrete nature. Most other features with significant t-test probabilities provided also a leave-one-out AUC value which was greater than the maximum chance criterion of $38/56 = 0.679$. The maximum chance criterion is given by the accuracy when all lesions are classified as belonging to the class with the most members.

Tables 7.9-7.11 list for completeness the class mean values, the class standard deviations, the t-test probabilities and the leave-one-out AUC values for the individual features.

In summary, registration and segmentation refinement had on average a positive effect on the classification performance of the individual features. The t-test probabilities were usually in accordance with the AUC classification performance of the individual features. Any differences are most likely due to the violations of the underlying t-test assumption of a binormal distribution. As expected, the least correlated features were not necessarily the most relevant features. The next section will show what effect this had when combining the least correlated features.
Figure 7.9: Influence of registration and segmentation refinement (x-axis, Table 7.6) on the probability that the mean feature value of benign and malignant lesions is equal. The average probability (y-axis) denotes the mean with respect to a set of features (legend).

Figure 7.10: Influence of registration and segmentation refinement (x-axis, Table 7.6) on the leave-one-out area under the ROC curve of individual features. The average area (y-axis) denotes the mean with respect to a set of features (legend).
### Table 7.7: T-test probability of individual features for 5 methods as described in Table 7.6. The list is sorted according to mean result of each row (last column). Probabilities that were significant at the 5% level are marked with a box. Features belonging to the set of the 11 least correlated features are marked with →.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Probability that class means are equal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Man.</td>
</tr>
<tr>
<td>$K_{quw}$</td>
<td>0.003</td>
</tr>
<tr>
<td>$K_{ser}$</td>
<td>0.007</td>
</tr>
<tr>
<td>$K_{mwo}$</td>
<td>0.008</td>
</tr>
<tr>
<td>$K_{rmwo}$</td>
<td>0.018</td>
</tr>
<tr>
<td>$K_{tmx}$</td>
<td>0.031</td>
</tr>
<tr>
<td>$M_{emb}$</td>
<td>0.028</td>
</tr>
<tr>
<td>$K_{rie}$</td>
<td>0.036</td>
</tr>
<tr>
<td>$T_{sue}$</td>
<td>0.127</td>
</tr>
<tr>
<td>$T_{ent}$</td>
<td>0.141</td>
</tr>
<tr>
<td>$T_{var}$</td>
<td>0.116</td>
</tr>
<tr>
<td>$S_{top}$</td>
<td>0.248</td>
</tr>
<tr>
<td>$T_{irmwv}$</td>
<td>0.115</td>
</tr>
<tr>
<td>$T_{asm}$</td>
<td>0.305</td>
</tr>
<tr>
<td>$T_{irmwv}$</td>
<td>0.262</td>
</tr>
<tr>
<td>$O_{volume}$</td>
<td>0.297</td>
</tr>
<tr>
<td>$K_{lpe}$</td>
<td>0.389</td>
</tr>
<tr>
<td>$K_{sde}$</td>
<td>0.390</td>
</tr>
<tr>
<td>$K_{slf}$</td>
<td>0.177</td>
</tr>
<tr>
<td>$K_{rme}$</td>
<td>0.536</td>
</tr>
<tr>
<td>$S_{rle}$</td>
<td>0.644</td>
</tr>
<tr>
<td>$S_{comp}$</td>
<td>0.646</td>
</tr>
<tr>
<td>$O_{age}$</td>
<td>0.568</td>
</tr>
<tr>
<td>$M_{mm gm}$</td>
<td>0.743</td>
</tr>
<tr>
<td>$S_{rrl}$</td>
<td>0.635</td>
</tr>
<tr>
<td>$S_{mrrgh}$</td>
<td>0.949</td>
</tr>
</tbody>
</table>
### Table 7.8: Leave-one-out area under the ROC curve (AUC) for individual features for 5 methods as described in Table 7.6. The list is sorted according to the mean AUC of each row (last column). Configurations which achieved t-test probabilities below 0.05 are marked with a box. Features belonging to the set of the 11 least correlated features are marked with →.

<table>
<thead>
<tr>
<th>Least correlated</th>
<th>Feature</th>
<th>Leave-one-out area under ROC curve (AUC)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Man.</td>
</tr>
<tr>
<td>→</td>
<td>$K_{ser}$</td>
<td>0.822</td>
</tr>
<tr>
<td></td>
<td>$K_{gwo}$</td>
<td>0.747</td>
</tr>
<tr>
<td>→</td>
<td>$K_{rie}$</td>
<td>0.699</td>
</tr>
<tr>
<td></td>
<td>$M_{emg}$</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td>$K_{rmwo}$</td>
<td>0.632</td>
</tr>
<tr>
<td></td>
<td>$T_{var}$</td>
<td>0.642</td>
</tr>
<tr>
<td>→</td>
<td>$K_{mwo}$</td>
<td>0.636</td>
</tr>
<tr>
<td>→</td>
<td>$T_{sue}$</td>
<td>0.624</td>
</tr>
<tr>
<td></td>
<td>$T_{ent}$</td>
<td>0.626</td>
</tr>
<tr>
<td></td>
<td>$T_{asm}$</td>
<td>0.618</td>
</tr>
<tr>
<td>→</td>
<td>$K_{slf}$</td>
<td>0.646</td>
</tr>
<tr>
<td></td>
<td>$K_{rme}$</td>
<td>0.619</td>
</tr>
<tr>
<td>→</td>
<td>$K_{sde}$</td>
<td>0.618</td>
</tr>
<tr>
<td></td>
<td>$K_{tmx}$</td>
<td>0.563</td>
</tr>
<tr>
<td>→</td>
<td>$K_{tpe}$</td>
<td>0.569</td>
</tr>
<tr>
<td>→</td>
<td>$T_{imuv}$</td>
<td>0.582</td>
</tr>
<tr>
<td></td>
<td>$S_{top}$</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td>$S_{ser}$</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>$O_{age}$</td>
<td>0.517</td>
</tr>
<tr>
<td>→</td>
<td>$O_{volume}$</td>
<td>0.513</td>
</tr>
<tr>
<td></td>
<td>$S_{rrl}$</td>
<td>0.531</td>
</tr>
<tr>
<td>→</td>
<td>$M_{mmmg}$</td>
<td>0.477</td>
</tr>
<tr>
<td>→</td>
<td>$S_{comp}$</td>
<td>0.462</td>
</tr>
<tr>
<td></td>
<td>$T_{irmuv}$</td>
<td>0.365</td>
</tr>
<tr>
<td></td>
<td>$S_{mrgrh}$</td>
<td>0.117</td>
</tr>
</tbody>
</table>
### Table 7.9: Characteristics of shape, margin and other features.

The P-value denotes the t-test probability, i.e., the probability that the mean values of benign and malignant lesions were the same. Statistically significant probabilities at the 5% level are marked with a box. AUC stands for the leave-one-out area under the ROC curve.

<table>
<thead>
<tr>
<th>Feature</th>
<th>Method</th>
<th>Mean value (± standard deviation)</th>
<th>P-value</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Benign lesions</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Malignant lesions</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_{comp}$</td>
<td>Man.</td>
<td>35.57 ± 15.68</td>
<td>0.646</td>
<td>0.462</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>40.19 ± 24.96</td>
<td>0.415</td>
<td>0.472</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>41.31 ± 27.57</td>
<td>0.377</td>
<td>0.490</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>40.75 ± 24.47</td>
<td>0.310</td>
<td>0.536</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>39.47 ± 23.95</td>
<td>0.946</td>
<td>0.158</td>
</tr>
<tr>
<td>$S_{rsl}$</td>
<td>Man.</td>
<td>0.19 ± 0.13</td>
<td>0.635</td>
<td>0.531</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>0.21 ± 0.13</td>
<td>0.947</td>
<td>0.080</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>0.19 ± 0.15</td>
<td>0.742</td>
<td>0.545</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>0.19 ± 0.15</td>
<td>0.715</td>
<td>0.552</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>0.20 ± 0.12</td>
<td>0.614</td>
<td>0.545</td>
</tr>
<tr>
<td>$S_{erl}$</td>
<td>Man.</td>
<td>2.03 ± 0.11</td>
<td>0.644</td>
<td>0.539</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>2.03 ± 0.08</td>
<td>0.476</td>
<td>0.534</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>2.03 ± 0.10</td>
<td>0.680</td>
<td>0.493</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>2.05 ± 0.10</td>
<td>0.257</td>
<td>0.583</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>2.03 ± 0.10</td>
<td>0.539</td>
<td>0.511</td>
</tr>
<tr>
<td>$S_{top}$</td>
<td>Man.</td>
<td>-10.94 ± 17.45</td>
<td>0.248</td>
<td>0.466</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>-52.39 ± 68.26</td>
<td>0.175</td>
<td>0.550</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>-49.89 ± 71.90</td>
<td>0.178</td>
<td>0.546</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>-53.56 ± 65.76</td>
<td>0.059</td>
<td>0.583</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>-58.11 ± 69.67</td>
<td>0.151</td>
<td>0.553</td>
</tr>
<tr>
<td>$S_{mervh}$</td>
<td>Man.</td>
<td>0.00 ± 0.00</td>
<td>0.949</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>0.00 ± 0.00</td>
<td>0.887</td>
<td>0.193</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>0.00 ± 0.00</td>
<td>0.506</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>0.00 ± 0.00</td>
<td>0.828</td>
<td>0.313</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>0.00 ± 0.00</td>
<td>0.904</td>
<td>0.311</td>
</tr>
<tr>
<td>$M_{emg}$</td>
<td>Man.</td>
<td>0.18 ± 0.06</td>
<td>0.028</td>
<td>0.662</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>0.18 ± 0.06</td>
<td>0.026</td>
<td>0.676</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>0.17 ± 0.06</td>
<td>0.099</td>
<td>0.713</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>0.17 ± 0.07</td>
<td>0.032</td>
<td>0.660</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>0.17 ± 0.06</td>
<td>0.033</td>
<td>0.682</td>
</tr>
<tr>
<td>$M_{vmemg}$</td>
<td>Man.</td>
<td>0.01 ± 0.00</td>
<td>0.743</td>
<td>0.477</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>0.01 ± 0.00</td>
<td>0.747</td>
<td>0.525</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>0.01 ± 0.00</td>
<td>0.944</td>
<td>0.194</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>0.01 ± 0.00</td>
<td>0.707</td>
<td>0.461</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>0.01 ± 0.00</td>
<td>0.375</td>
<td>0.533</td>
</tr>
<tr>
<td>$O_{age}$</td>
<td>Man.</td>
<td>46.16 ± 9.84</td>
<td>0.568</td>
<td>0.517</td>
</tr>
<tr>
<td></td>
<td>Unreg.</td>
<td>47.83 ± 10.30</td>
<td>0.297</td>
<td>0.513</td>
</tr>
<tr>
<td></td>
<td>Rigid</td>
<td>2912.29 ± 10.30</td>
<td>0.291</td>
<td>0.516</td>
</tr>
<tr>
<td></td>
<td>NR20V</td>
<td>2694.08 ± 2897.00</td>
<td>0.292</td>
<td>0.515</td>
</tr>
<tr>
<td></td>
<td>NR10</td>
<td>2691.11 ± 2899.50</td>
<td>0.282</td>
<td>0.518</td>
</tr>
</tbody>
</table>

Statistically significant probabilities at the 5% level are marked with a box. AUC stands for the leave-one-out area under the ROC curve.
<table>
<thead>
<tr>
<th>Feature Method</th>
<th>Mean value (± standard deviation)</th>
<th>P-value</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K_{rme}$ Man.</td>
<td>2.13 ± 1.42</td>
<td>2.32 ± 0.83</td>
<td>0.536</td>
</tr>
<tr>
<td>Unreg.</td>
<td>2.13 ± 1.37</td>
<td>2.29 ± 0.78</td>
<td>0.587</td>
</tr>
<tr>
<td>Rigid</td>
<td>2.11 ± 1.40</td>
<td>2.35 ± 0.79</td>
<td>0.421</td>
</tr>
<tr>
<td>NR20V</td>
<td>2.10 ± 1.42</td>
<td>2.32 ± 0.79</td>
<td>0.459</td>
</tr>
<tr>
<td>NR10</td>
<td>1.90 ± 1.22</td>
<td>2.06 ± 0.77</td>
<td>0.539</td>
</tr>
<tr>
<td>$K_{slf}$ Man.</td>
<td>3.21 ± 2.07</td>
<td>2.41 ± 2.05</td>
<td>0.177</td>
</tr>
<tr>
<td>Unreg.</td>
<td>3.09 ± 1.60</td>
<td>2.56 ± 2.23</td>
<td>0.374</td>
</tr>
<tr>
<td>Rigid</td>
<td>3.11 ± 1.61</td>
<td>2.56 ± 2.16</td>
<td>0.343</td>
</tr>
<tr>
<td>NR20V</td>
<td>2.99 ± 1.49</td>
<td>2.63 ± 2.08</td>
<td>0.517</td>
</tr>
<tr>
<td>NR10</td>
<td>2.86 ± 1.10</td>
<td>2.51 ± 1.92</td>
<td>0.480</td>
</tr>
<tr>
<td>$K_{rmx}$ Man.</td>
<td>340.00 ± 104.94</td>
<td>272.37 ± 107.69</td>
<td>0.031</td>
</tr>
<tr>
<td>Unreg.</td>
<td>345.00 ± 98.83</td>
<td>272.37 ± 115.53</td>
<td>0.026</td>
</tr>
<tr>
<td>Rigid</td>
<td>365.00 ± 84.38</td>
<td>284.21 ± 122.95</td>
<td>0.015</td>
</tr>
<tr>
<td>NR20V</td>
<td>365.00 ± 84.38</td>
<td>284.21 ± 128.18</td>
<td>0.018</td>
</tr>
<tr>
<td>NR10</td>
<td>370.00 ± 81.02</td>
<td>277.11 ± 120.90</td>
<td>0.005</td>
</tr>
<tr>
<td>$K_{sde}$ Man.</td>
<td>55.79 ± 43.39</td>
<td>64.56 ± 31.07</td>
<td>0.390</td>
</tr>
<tr>
<td>Unreg.</td>
<td>55.51 ± 43.60</td>
<td>64.35 ± 31.12</td>
<td>0.388</td>
</tr>
<tr>
<td>Rigid</td>
<td>55.60 ± 44.36</td>
<td>65.28 ± 31.15</td>
<td>0.349</td>
</tr>
<tr>
<td>NR20V</td>
<td>55.25 ± 44.41</td>
<td>64.51 ± 31.10</td>
<td>0.371</td>
</tr>
<tr>
<td>NR10</td>
<td>51.26 ± 40.48</td>
<td>60.06 ± 30.50</td>
<td>0.370</td>
</tr>
<tr>
<td>$K_{rie}$ Man.</td>
<td>1.11 ± 0.97</td>
<td>1.68 ± 0.91</td>
<td>0.036</td>
</tr>
<tr>
<td>Unreg.</td>
<td>1.09 ± 0.93</td>
<td>1.63 ± 0.85</td>
<td>0.036</td>
</tr>
<tr>
<td>Rigid</td>
<td>1.06 ± 0.93</td>
<td>1.68 ± 0.86</td>
<td>0.018</td>
</tr>
<tr>
<td>NR20V</td>
<td>1.06 ± 0.94</td>
<td>1.68 ± 0.86</td>
<td>0.017</td>
</tr>
<tr>
<td>NR10</td>
<td>1.01 ± 0.91</td>
<td>1.55 ± 0.83</td>
<td>0.033</td>
</tr>
<tr>
<td>$K_{mwo}$ Man.</td>
<td>0.02 ± 0.03</td>
<td>0.06 ± 0.06</td>
<td>0.008</td>
</tr>
<tr>
<td>Unreg.</td>
<td>0.01 ± 0.02</td>
<td>0.06 ± 0.06</td>
<td>0.007</td>
</tr>
<tr>
<td>Rigid</td>
<td>0.01 ± 0.02</td>
<td>0.06 ± 0.07</td>
<td>0.003</td>
</tr>
<tr>
<td>NR20V</td>
<td>0.01 ± 0.02</td>
<td>0.06 ± 0.07</td>
<td>0.005</td>
</tr>
<tr>
<td>NR10</td>
<td>0.01 ± 0.01</td>
<td>0.05 ± 0.06</td>
<td>0.003</td>
</tr>
<tr>
<td>$K_{gwo}$ Man.</td>
<td>0.96 ± 0.58</td>
<td>0.41 ± 0.63</td>
<td>0.003</td>
</tr>
<tr>
<td>Unreg.</td>
<td>0.99 ± 0.56</td>
<td>0.45 ± 0.61</td>
<td>0.002</td>
</tr>
<tr>
<td>Rigid</td>
<td>1.01 ± 0.56</td>
<td>0.46 ± 0.63</td>
<td>0.002</td>
</tr>
<tr>
<td>NR20V</td>
<td>1.01 ± 0.57</td>
<td>0.43 ± 0.62</td>
<td>0.001</td>
</tr>
<tr>
<td>NR10</td>
<td>0.86 ± 0.43</td>
<td>0.34 ± 0.51</td>
<td>0.000</td>
</tr>
<tr>
<td>$K_{rmwo}$ Man.</td>
<td>1.71 ± 3.40</td>
<td>5.23 ± 5.64</td>
<td>0.018</td>
</tr>
<tr>
<td>Unreg.</td>
<td>1.04 ± 1.85</td>
<td>4.70 ± 5.16</td>
<td>0.005</td>
</tr>
<tr>
<td>Rigid</td>
<td>0.62 ± 1.37</td>
<td>5.04 ± 5.25</td>
<td>0.001</td>
</tr>
<tr>
<td>NR20V</td>
<td>0.70 ± 1.54</td>
<td>4.89 ± 5.36</td>
<td>0.002</td>
</tr>
<tr>
<td>NR10</td>
<td>0.53 ± 0.87</td>
<td>4.52 ± 5.20</td>
<td>0.002</td>
</tr>
</tbody>
</table>

Table 7.10: Characteristics of kinetic features. The P-value denotes the t-test probability, i.e. the probability that the mean values of benign and malignant lesions were the same. Statistically significant probabilities at the 5% level are marked with a [box]. AUC stands for the leave-one-out area under the ROC curve.
Table 7.11: Characteristics of kinetic features, texture features and radiological scores. The P-value denotes the t-test probability, i.e. the probability that the mean values of benign and malignant lesions were the same. Statistically significant probabilities at the 5% level are marked with a box. AUC stands for the leave-one-out area under the ROC curve.
7.3 Results

7.3.5 Multivariate Classification

This section describes the classification performance when combining features and compares it with the performance achieved by the cumulative scores determined by radiologists and by individual features. Features were combined with stepwise multivariate logistic regression analysis using a 20% forward selection and 25% backward elimination criterion. The classification performance was assessed based on the area under the ROC curve (AUC) created from leave-one-out tests.

Table 7.12 shows the classification result after combining the 11 least correlated features for 6 feature extraction methods. The maximum washout ($K_{mwo}$) was selected as first features by all image based classifiers, followed by either the signal enhancing ratio ($K_{ser}$) or the late peripheral enhancement ($K_{lpe}$). Two additional features were included by the classifier based on manual segmentations. A clear preference for kinetic features can be observed.

Figure 7.11 depicts the corresponding empirical ROC curves. The curves were distributed over two graphs to ease visualization. The result for the unregistered images was displayed on both graphs to help comparison. The top graph shows that image based classifier, i.e. manual segmentation and no registration, were better than classifiers based on the radiological score or on random features. An improvement of the image based classifiers due to registration can be observed in the bottom graph. These differences can be better appreciated by looking at the smooth ROC curves in Figure 7.12.

The area under the smooth ROC curve was statistically significantly greater for classifiers which were based on registered images than for classifiers based on random features or on the

<table>
<thead>
<tr>
<th>Method</th>
<th>Included features</th>
<th>Sens</th>
<th>Spec</th>
<th>Acc</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
<td>3rd</td>
<td>4th</td>
<td>5th</td>
</tr>
<tr>
<td>Man.</td>
<td>$K_{mwo}$</td>
<td>$K_{ser}$</td>
<td>$K_{lpe}$</td>
<td>$T_{imu}$</td>
<td>$K_{srf}$</td>
</tr>
<tr>
<td>Unreg.</td>
<td>$K_{mwo}$</td>
<td>$K_{ser}$</td>
<td>$K_{lpe}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigid</td>
<td>$K_{mwo}$</td>
<td>$K_{lpe}$</td>
<td>$K_{ser}$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR20V</td>
<td>$K_{mwo}$</td>
<td>$K_{ser}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR10</td>
<td>$K_{mwo}$</td>
<td>$K_{ser}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>$R_6$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.12: Leave-one-out classification results for 6 feature extraction methods (Table 7.6), where the 11 least correlated features (Table 7.5) were combined using stepwise multivariate logistic regression. Sensitivity (Sens), specificity (Spec) and accuracy (Acc) refer to the values which were attained when maximizing accuracy. AUC denotes the area under the empirical or the smooth ROC curve.
Figure 7.11: Empirical ROC curves for the radiological scores and the 11 least correlated features (Table 7.5) for the methods described in Table 7.6. The bottom graph shows the impact that the registration had on the empirical ROC curve.
Figure 7.12: Smooth ROC curves for the radiological scores and the 11 least correlated features (Table 7.5) for the methods described in Table 7.6. The curves were generated by fitting binomial distributions to the data using the ROCKIT program from [Metz et al., 1998].
first radiological score, see Table 7.13. The classifier based on rigidly registered images was also statistically significantly better than the classifier based on the original image sequence (Unreg.). Note that the parameter fitting varies slightly, depending on which two datasets are compared. The stated AUC values represent the values when methods were compared to Unreg.

The sensitivity of the classification result to the initial crude outline was tested by randomly changing the volume of the outline by either -33%, -20%, 0%, 25% or 50% before segmentation refinement. Comparing Table 7.14 with Table 7.12, it can be observed that very similar empirical AUC values were achieved although different features were selected.

The changes in feature values due to rigid registration can be seen in Figure 7.13 for the first two features included in the classifiers. Improvements as well as worsenings can be observed. On average, the feature values were slightly more different for benign and

<table>
<thead>
<tr>
<th>Method</th>
<th>Score1</th>
<th>Score2</th>
<th>Man.</th>
<th>Unreg.</th>
<th>Rigid</th>
<th>NR20V</th>
<th>NR10</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUC</td>
<td>0.685</td>
<td>0.753</td>
<td>0.786</td>
<td>0.810</td>
<td>0.862</td>
<td>0.859</td>
<td>0.833</td>
</tr>
<tr>
<td>Random</td>
<td>0.686</td>
<td>0.481</td>
<td>0.239</td>
<td>0.159</td>
<td>0.088</td>
<td>0.027</td>
<td>0.020</td>
</tr>
<tr>
<td>Score1</td>
<td>0.685</td>
<td>0.174</td>
<td>0.109</td>
<td>0.058</td>
<td>0.019</td>
<td>0.011</td>
<td>0.031</td>
</tr>
<tr>
<td>Score2</td>
<td>0.753</td>
<td>0.328</td>
<td>0.206</td>
<td>0.082</td>
<td>0.054</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>Man.</td>
<td>0.786</td>
<td>0.294</td>
<td>0.114</td>
<td>0.072</td>
<td>0.213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unreg.</td>
<td>0.810</td>
<td>0.294</td>
<td>0.114</td>
<td>0.072</td>
<td>0.213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rigid</td>
<td>0.862</td>
<td>0.419</td>
<td>0.286</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NR20V</td>
<td>0.859</td>
<td>0.419</td>
<td>0.286</td>
<td>0.104</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.13: Probabilities that the area under smooth ROC curve (AUC) is the same for two methods. Statistically significant results at the 5% level are marked by a box. The methods are described in Table 7.6.

<table>
<thead>
<tr>
<th>Method</th>
<th>Included features</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
<th>4th</th>
<th>Sens</th>
<th>Spec</th>
<th>Acc</th>
<th>AUC</th>
<th>empirical</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unreg.</td>
<td>$K_{ser}$ $K_{mwo}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>79</td>
<td>72</td>
<td>77</td>
<td>0.806</td>
<td></td>
</tr>
<tr>
<td>Rigid</td>
<td>$K_{mwo}$ $K_{ser}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82</td>
<td>83</td>
<td>82</td>
<td>0.850</td>
<td></td>
</tr>
<tr>
<td>NR20V</td>
<td>$K_{mwo}$ $K_{ser}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>82</td>
<td>78</td>
<td>80</td>
<td>0.843</td>
<td></td>
</tr>
<tr>
<td>NR10</td>
<td>$K_{mwo}$ $K_{sfl}$ $K_{ser}$ $K_{lpe}$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>89</td>
<td>56</td>
<td>79</td>
<td>0.844</td>
<td></td>
</tr>
</tbody>
</table>

Table 7.14: Leave-one-out classification result when randomly changing the volume of the initial crude outline by either -33%, -20%, 0%, 25% or 50% before segmentation refinement.
7.3 Results

Figure 7.13: Change of feature values due to rigid registration for (left) the maximum washout \( (K_{mwo}) \) and (right) the signal enhancing ratio \( (K_{ser}) \). The value pair \((-4.94, -5.10)\) for a malignant lesion was omitted from the right figure to improve visualization.

malignant lesion after registration.

The mean of the maximum washout \( (K_{mwo}) \) decreased from 1.2% to 0.9% for benign lesions and increased from 5.6% to 5.7% for malignant lesions after rigid registration. Example slices of the lesion which had the largest \( K_{mwo} \) decrease are shown in Figure 7.14. A reduction in motion artifacts and a more similar lesion appearance for the whole image sequence can be observed. \( K_{mwo} \) changed from 3% to 0% making this benign lesion more similar to other benign lesions. This improved the classification performance because the order of benign and malignant lesion, when ranking \( K_{mwo} \), was changed to the better.

The mean signal enhancing ratio \( (K_{ser}) \) increased from 1.61 to 1.62 for benign lesions and decreased from 0.42 to 0.41 for malignant lesions. Example slices of the lesion which had the largest \( K_{ser} \) increase are shown in Figure 7.15. This malignant lesion has a typical benign appearance with a late enhancement and no apparent washout. Its \( K_{ser} \) value for the original images is 3.14, which is very high for a malignant lesion. Note that \( K_{ser} = K_{gwo}/K_{rie} \), where \( K_{gwo} \) denotes the general washout and \( K_{rie} \) the relative initial enhancement. Rigid registration further increased this value to 3.65. This was caused by a decrease in \( K_{rie} \) from 0.46 to 0.39, which is indistinguishable in Figure 7.15. The classification performance was not affected by this change, since the order of benign and malignant lesions was maintained.

It can be observed by comparing Table 7.12 with Table 7.8, that the best individual feature \( K_{ser} \) was in most cases better than the combined features. This was because classifiers were optimized with respect to the likelihood ratio and not the area under the ROC curve.
7.4 Summary

A computer aided diagnosis system was created which achieved an area under the ROC curve of 0.86. Classification of MR breast lesion based on stepwise linear discriminant or logistic regression analysis of extracted features from lesion segmentations has been reported previously, see Table 4.1. These studies accomplished areas under the ROC curve between 0.80 and 0.96. The classification performance of this work is also within this range. Of the three studies which had a better classification performance [Gilhuijs et al., 1998; Gilhuijs et al., 2002; Gibbs and Turnbull, 2003], only one [Gilhuijs et al., 2002] complied with the 5 observation per feature recommendation.

Refining crude lesion outlines provided visually satisfactory results, allowed segmentation propagation after registration and led together with registration to classification improvements. The classification performance was robust to volume changes of the initial crude outlines. Deriving the crude outlines from the manual segmentation could have biased the outcome. The segmentation refinement for the CAD system is however envisaged as an in-
teractive tool, rather than a retrospective one. Over- or underestimation of the lesion size could therefore easily be corrected by the user.

Image registration and segmentation refinement generally improved the classification performance. The improvement in classification performance due to image registration was statistically significant for rigid registration, and was close to being statistically significant for 20mm non-rigid registration. More data is required for reliably detecting differences between these registration methods because only a few images had severe motion artifacts. For example, a dataset of 201 benign and 201 malignant lesions would be required in order to detect, with a probability of 80%, a statistical significant difference at the 5% level between an AUC of 0.800 and an AUC of 0.875 [Hanley and McNeil, 1982].

The cumulative scores from the radiologists performed worse than the classifier based on image features. An improvement can be expected when the individual scores are optimized in a similar way to the other classifiers. No radiological assessment independent from the scoring scheme was available for comparison.

The classification results were assessed for bias by comparing them with the results of similarly trained classifiers based on random features. Classifiers based on 11 image features derived after registration achieved statistically significantly better results than the classifier based on 11 random features. These improvements suggest that these results were not artifacts of the feature selection process or the leave-one-out test procedure. However, for 25 feature candidates both random features and image features derived after registration achieved a high performance (AUC∈[0.87,0.92]). This indicates that the better classification performance for the 25 feature candidates was caused by the feature selection process or the leave-one-out test procedure rather than the feature properties.

Less than half of the promising features provided statistically significantly differences in the mean value of benign and malignant lesions. The classification performance was generally best for kinetic features and worst for shape features. Excluded features are not necessarily useless, since they might have been excluded because of their high correlation to the included features.

There was no advantage in using logistic regression analysis instead of linear discriminant analysis, as both providing very similar classification performances.

In conclusion, a successful computer aided diagnosis system has been created. The inclusion of image registration into this CAD system was beneficial as it led to a statistically significant increase of the area under the ROC curve from 0.81 to 0.86.
Chapter 8

Conclusions and Future Work

The first objective of this thesis was to develop a technique for the validation of registration algorithms for correcting for patient motion in DCE MR mammography. The second objective was to improve the classification of breast lesions by creating a computer aided diagnosis (CAD) system for DCE MR mammography which includes image registration. This final chapter summarizes how this was achieved, highlights the novel contributions and also suggests future work.

8.1 Summary

The need for aligning the image sequence obtained during DCE MR mammography to improve MR breast lesion classification was identified in Chapter 1. Previous visual assessment of an existing registration algorithm [Rueckert et al., 1999b] had also shown that it significantly improved the quality of the subtraction images [Denton et al., 1999]. Early work undertaken for this thesis revealed, however, that this registration algorithm changed the volume of the breast lesions by an unacceptable amount [Tanner et al., 2000]. This was the first study which recognized this problem and recommended constraint of DCE MR mammography registration to preserve volume.

8.1.1 Biomechanical Breast Models

The literature review of registration algorithms for DCE MR mammography in Chapter 2 concluded that their assessment was inadequate. A novel validation method was therefore proposed based on simulating plausible deformations using finite element methods. In the case of DCE MR mammography, the task was to create biomechanical breast models which can produce deformations that are likely to occur during the acquisition of the image sequence. Chapter 3 reviewed continuum mechanics, material models, finite element methods, mechanical breast properties and existing biomechanical breast models. The reported elastic values for the breast tissue types were inconsistent and the evaluation of the published biomechanical breast models was found to be insufficient. Biomechanical breast models were therefore created for a wide range of material properties and their accuracies of predicting the displacement of internal breast structures were compared for two datasets (Chapter 5). Biomechanical breast models mainly vary with respect to the mesh generation, the bound-
ary conditions employed, the tissue properties assumed and the solution strategies applied. The influence of all of these aspects on the accuracy was assessed. Experiments included a comparison of transverse isotropic material models, Mooney-Rivlin hyperelastic models and finite deformation formulations with simpler approaches. These comparisons and the use of transverse isotropic material models for predicting breast deformations are novel. Models with appropriate Poisson’s ratios and suitable elastic properties reduced the mean (maximum) displacement error from 6.6mm (12.4mm) to 2.1mm (3.4mm) for accurate boundary conditions. This compares favourably with results reported by other groups. A linear model with a Young’s modulus of 1kPa for fat and of 1.5kPa for fibroglandular tissue was identified as a suitable model for the validation of the registration.

8.1.2 Validation of Registration

Chapter 6 described the validation of the non-rigid registration algorithm for DCE MR mammography. Ten pre- and post-contrast image pairs which had no visible motion artifacts were selected from a large database. Plausible deformations were created for 6 deformation scenarios and 2 deformation magnitudes. Normalized mutual information yielded the best results when comparing the performance for 6 similarity measures. Constraining non-rigid registration to preserve volume, as suggested by [Rohlffing and Maurer, 2001], further improved the registration accuracy. This was especially valuable in the case of the most flexible deformation models and for reducing the error within the lesion. The better registration strategies for the training set were 40mm single-resolution FFD registration with $\mu \in [0,0.8]$ and 20mm single-resolution FFD registration with $\mu \in [0.7,0.8]$, where $\mu$ is the weight of the volume preservation regularization term in equation (6.1) on page 183. These strategies reduced on average the 95th percentile target registration error for the test set from 3.00mm to 1.10mm for the whole breast tissue and from 1.63mm to 0.46mm for the enhancing lesion. Affine registration reduced these values to 1.27mm and 0.57mm, respectively. The DCE MR mammograms had in contrast a mean voxel dimension of $1.36 \times 1.36 \times 3.18 \text{mm}^3$. A comparison with other registration algorithms is not possible, because their target registration error for DCE mammography is unknown. The difference in intensity due to a contrast agent was identified as the main cause for the volume change of the unconstrained registration algorithm.

8.1.3 CAD System

After having determined the better registration configurations in Chapter 6, the next task was to design a CAD system for DCE MR mammography. A suitable framework for the CAD system was identified by reviewing the literature (Chapter 4). It consisted of semi-automatic
lesion segmentation, extraction of promising features, selection of the least correlated features, classification based on a logistic regression model and evaluation by leave-one-out receiver operator characteristics (ROC) analysis. Chapter 7 described the design and the evaluation of this CAD system. A new semi-automatic segmentation method was specially devised to reduce the workload of the radiologists. It is based on a constrained maximum a posteriori approach, it refines crude outlines, and it was used to propagate segmentations to image sequences registered in a different way. The best segmentation refinement configuration was selected based on visual assessment. The CAD system achieved a leave-one-out area under the ROC curve of 0.86, which is within the range of reported classification performances. The conservative use of features ensured, in comparison to other studies, that this performance was not an artifact of the feature selection process or the leave-one-out test procedure. The classification results were robust with respect to volume changes of the initial crude outlines. Worse results were obtained without segmentation refinement and image registration. Results from rigidly registered images were statistically significantly better at the 5% level than results from the original images.

8.2 Conclusions and Discussion

This thesis showed that incorporating image registration into a CAD system for DCE MR mammography is beneficial. The classification performance for rigidly registered images was statistically significantly better than for the original images. Volume-preserving non-rigid registration led to very similar classification results as rigid registration. The small dataset prohibited reliable detection of a statistically significant difference for the non-rigid registration configurations, because only a few images were severely misaligned within the region of the lesion. The developed CAD system achieved, with an area under the ROC curve of 0.86 for leave-on-out tests, a state of the art classification performance, which was proven not to be the artifact of the feature selection process or the leave-one-out test procedure. The study confirmed that too many feature candidates can lead to spuriously good classification results for leave-one-out tests, which is often overlooked. The devised semi-automatic segmentation method provided slightly better classification results than the manual segmentation and has great potential as an interactive tool.

The novel method of assessing the registration accuracy by simulating biomechanically plausible deformations enabled the selection of a registration configuration for the CAD system. Constraining non-rigid registration to preserve volume increased its registration accuracy substantially within the region of the lesion. The better registration configura-
tions aligned the enhancing lesion with an accuracy of 0.32mm on average, which cannot be compared with other studies since their accuracy is unknown. A possible shortcoming of the assessment was that results could have been biased because the simulated breast deformations were more homogeneous and continuous than real deformations. Nevertheless, constraining non-rigid registrations to preserve volume to some extent is likely to be advantageous for aligning other contrast-enhanced images which captured volume preserving deformations.

Biomechanical breast models, with suitable configuration were able to predict the displacement of internal breast structure with a mean accuracy of 2.1mm when breasts were compressed by 20% and displacement boundary conditions were applied. This is better than anything that has been reported by other groups to date. Neither finite strain formulations and hyperelastic models nor transverse isotropic models improved this accuracy. This study was limited by the small number of datasets, which prohibits any prediction for the general population. Yet the findings suggest that linear biomechanical breast models may be a viable alternative to non-linear and finite strain formulations when accurate displacement boundary conditions and suitable patient specific model parameters are available.

8.3 Future Work

8.3.1 CAD System

While the CAD system has shown promising results, more datasets are required for reliably detecting significant differences between registration methods. Improved results are likely to arise with more observations, since more features can be assessed when building the classifier.

It is also recommended that the semi-automatic segmentation method is developed into an interactive segmentation tool. In this way, radiologists are able to extract a lesion which is most consistent with their initial outline for the whole image sequence while being able to control its size. In this form, it might be enough to provide the outline of just one slice.

The CAD system can be made more informative by augmenting its classification suggestions by confidence measures. Such a measure could reflect the probability of malignancy as estimated from the pathology results of the N most similar cases. Another measure could indicate to what extent the case is novel, based on the estimated probability density of the training data [Markou and Singh, 2003]. For better acceptance by the radiologist, an explanation could be provided for the classification suggestions. This could come in the form of depicting the most similar lesions with available pathology [Doi, 2004], of extracting classification rules or by visualizing the decision boundaries [Andrews et al., 1995].

Ultimately, a prospective clinical trial should be performed to evaluate the clinical benefits
of the classification suggestions from the CAD system to aid the decision making of the radiologist.

\subsection*{8.3.2 Biomechanical Breast Models}

The biomechanical breast models could be improved by incorporating sliding tissue, and if Cooper’s ligaments were visible on the images or could be inferred from anatomical knowledge and hence their local anisotropic influence could be modelled. Substantial improvements would, however, require knowledge of the patient-specific spatial distribution of the in-vivo elastic properties for the applied strain range. Estimation of these properties are especially difficult for high strains due to the non-linearities introduced.

The currently achieved mean (maximum) prediction error of 2.1mm (3.3mm) is however already encouraging for guiding breast surgeons in cases where their only guidance is otherwise a wire and X-ray mammograms from the compressed breast. A system can be envisaged based on pre-operative DCE MR mammograms, a pre-operative patient-specific biomechanical breast model and registration of the breast surface extracted from the DCE MR mammogram to the breast surface captured during surgery. The main challenges will be to obtain accurate boundary conditions especially for the posterior breast boundary and after incision. Work is underway to develop such a system in our laboratory.

Biomechanical breast models have various other applications. They can for example provide prior information for ill-posed breast registration problems like X-ray mammography to DCE MR mammography registration. They can help to speed up registration by reducing the number of transformation parameters which need to be considered. They can be used for generating gold standard deformations for validation purposes of other breast registration tasks. Together with image registration, they can be employed to estimate the most plausible elastic properties \cite{Miga2002}.

\subsection*{8.3.3 Validation of Registration}

If the accuracy of further biomechanical breast models could be substantially improved, then these models should of course be employed to simulate better gold standard breast deformations for the validation of registration algorithms.

The current dataset of simulated breast deformations with realistic contrast changes constitutes, however, already a valuable test set for the registration community. Collaborations with other groups on optimizing and validating approaches for the intra-visit registration of DCE MR breast images are welcomed.
Bibliography


Bibliography 248


Bibliography


Publications

Articles in Journals


Articles in Conference Proceedings


