A medical image-based computer-aided diagnosis system for musculoskeletal disease and disorder.

Chuah, Tong Kuan

2012

Chuah, T. K. (2012). A medical image-based computer-aided diagnosis system for musculoskeletal disease and disorder. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/52048>

<https://doi.org/10.32657/10356/52048>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**A MEDICAL IMAGE-BASED COMPUTER-AIDED
DIAGNOSIS SYSTEM FOR MUSCULOSKELETAL
DISEASE AND DISORDER**

CHUAH TONG KUAN
SCHOOL OF CHEMICAL AND BIOMEDICAL
ENGINEERING

2012

**A MEDICAL IMAGE-BASED COMPUTER-AIDED
DIAGNOSIS SYSTEM FOR MUSCULOSKELETAL
DISEASE AND DISORDER**

CHUAH TONG KUAN

School of Chemical and Biomedical Engineering

**A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirement for the degree of
Doctor of Philosophy**

Abstract

Musculoskeletal diseases such as osteoarthritis (OA) are responsible for a large number of disabilities among the world's population. With the increase of aging population and increase in the amount and complexity of medical data that clinicians need to handle, computer-aided diagnosis (CAD) system is becoming increasingly important in improving the efficiency, reproducibility and accuracy of the diagnosis process. CAD systems for breast cancer detection are currently being used in clinical practice but the use of CAD for musculoskeletal diseases is still under research. This thesis makes several advancements in the research and development of an image-based CAD system for musculoskeletal diseases, mainly focusing on OA and various aspects of the CAD system. The CAD system is designed to provide supplementary information to support medical decision, or to provide second opinion to the practitioner in cases of ambiguity.

An image-based CAD system typically has components such as a segmentation (or feature extraction) module, a measurement module, a classification module and/or a visualization module. This thesis made advancements in these components. Cartilage defect is an important biomarker for OA. It is important to be able to visualize damaged cartilage during diagnosis using medical images to ascertain the size and locations of the defects. To aid the visualization of damaged cartilage, the thesis first developed a visualization framework for visualizing cartilage damage or lesion in proton density weighted MR images. Using the cartilage visualization framework developed, it is possible to effectively display damaged cartilage. A metric has also been studied for its

ability to correlate with the percentage of damaged cartilage. A linear relationship between percentage of damaged cartilage and the metric studied was found.

As part of the advancement to classifying subjects with BML, the thesis investigated textural parameters as potential biomarker in separating between bone marrow with and without bone marrow lesion. Through statistical analysis, a set of parameters was identified and was further used for classification of image slices and subjects so that a second opinion could be provided by the computer. The classification results confirmed that image textural information of bone marrow can provide reasonably accurate results in differentiating between subjects with and without BML: the area under receiver operating characteristic (ROC) curve achieved is 0.914.

Having established the ability to classify subjects with and without BML, the thesis then deals with the development of an automated segmentation algorithm for the bone in knee MRI, by proposing a new termination criterion and an initialization strategy for active contours. Segmentation is an important and necessary step before features can be quantified to be used in the analysis and classification system. For automatic implementation of active contours in which different shapes need to be segmented, the proposed termination criterion demonstrated almost 50% and 60% total time reduction while achieving similar accuracy as compared with conventional pixel movement-based method in the segmentation of synthetic and real medical images, respectively. The initialization strategy worked as expected and achieved DSC of $96.7 \pm 1.1\%$ for the data validated.

Overall, the thesis has made a balanced advancement to various components of the CAD system for musculoskeletal diseases, forming the foundation for future work to incorporate more anatomical structures of the joints (ligaments, muscles etc.) and biomarkers in the diagnosis, and further improving the segmentation process.

Acknowledgements

This thesis work would not have been possible without the support of the people and organizations I would like to acknowledge here. First and foremost, I would like to thank my supervisor, Assistant Professor Poh Chueh Loo, who enthusiastically guided me and worked with me in all my learning and scientific researches. With his patient guidance and advice, I was challenged to perform better than I could be. He enlightened me to present my work in international conferences as well as scientific journals and I was motivated to do so. He encouraged me whenever there is hindrance in either research or publication. Overall, working with Professor Poh has not only given me the knowledge in scientific area but also it has made me into a better researcher with the required temperament, patience and confidence.

I would like to thank Nanyang Technological University for providing me financial support through Nanyang President Scholarship. It provided me with the needs in everyday life and campus housing. The working experience would not have been so rich and enjoyable without my colleagues Eric Van Reeth, Yu Peicong, Darwin Tay, and Yang Yuxin and Ho Jiahui. They were very helpful to have discussed with me when I have questions or doubts. Other than reading scientific publications, much of my knowledge and foundation of the field started from some simple discussions with them. I would also like to thank our collaborating radiologist Kenneth Seah and surgeon Denny Lie, who review papers, provide data and validations and provide research ideas for me.

My writing would not have been improved without the kind words and guidance of some anonymous reviewers. Although I do not know their names to thank them personally, I would be delighted to let them know that their help in reviewing my paper for me, has made me a better researcher, more focused on important issues, more meticulous to details and more considerate to readers.

I would also like to thank the supervisors who guided me when I was a teaching assistant: Assistant Professor Chew Sing Yian who spent time making sure with me that I trained myself enough for the teaching session, and Professor William Chen Wei Ning who has been friendly and supportive. Both of them have helped me to achieve the award of Best Teach Assistant for two consecutive semesters.

I would not forget the help I got from all the undergraduates or attachment students who worked with me. They include Lim Jun Hong, Leow Jiamin, Ong Yibin, Lim Cheng Wei, Abhirami Krishnaswamy, Jiang Yuan and Ong Eng Hui. Jun Hong has been helpful in running programs, doing manual segmentation and reporting results efficiently. All of them have helped me in one way or another, such as stimulating my thinking, decreasing my workload, and make me more familiar with things I have known.

Last but not least I would like to thank my mother for her bringing me up and everything she has given to me; and my sisters who have been encouraging all the time. I would like to thank my wonderful girlfriend who has been bringing me much joy at difficult times.

Table of Contents

Abstract	3
Acknowledgements	6
Table of Contents	8
List of Figures and Tables	11
Statement of Originality	13

Chapter 1

Introduction and Literature Review	15
1.1. Motivation	15
1.2. Diagnosis of Osteoarthritis	19
1.2.1. Symptoms and Biological Changes	19
System Level	20
Tissue Level	20
Cellular and Molecular Level	21
1.2.2. Diagnosis of Early OA using Magnetic Resonance Imaging	22
Morphological MRI	24
Quantitative MRI of the Cartilage Composition	25
MRI of Other Tissues	28
1.3. Overview of a Computer-aided Diagnosis System for musculoskeletal diseases ..	29
1.4. Image Understanding of Knee MRI	32
1.4.1. Medical Image Segmentation and Registration	33
Component Algorithms for MRI Knee Segmentation	36
Active Contour Segmentation	37
1.4.2. Feature Analysis of Anatomical Structures	40
1.5. Specific Objectives	42

Chapter 2

Visualization and Image Analysis of Damaged Cartilage using Knee Joint MRI	45
Introduction	45
2.1. Visualization of Damaged Cartilage of the Knee Joint Using Magnetic Resonance Imaging	46
2.1.1. Methods	48
Pre-processing and Segmentation of Cartilage	49
Determination of Threshold Value using Gaussian Curve Fitting	50
Visualization of the Damaged Portion of the Cartilage	52
Analysis	54
2.1.2. Results	54
2.1.3. Discussions and Conclusion	60
2.2. Characterizing Healthy and Damaged Cartilage in PD Images Using Histogram-based Parameters	64
2.2.1. Methods	65
Pre-processing and Segmentation of Cartilage	65

Analysis.....	65
2.2.2. Results.....	66
2.2.3. Discussions and Conclusion	69
Chapter Summary	70

Chapter 3

Texture Analysis and Classification of Bone Marrow Lesion in Knee MRI.....	72
Introduction.....	72
3.1. Texture Analysis of Bone Marrow Lesion in Knee MRI	73
3.1.1. Methods.....	75
MRI Data	75
Preprocessing of Images	76
Segmentation of Weight-bearing Region of Distal Femur – The Region of Interest	77
Grouping and Characterization of Images	79
Texture Analysis and Preliminary Feature Selection.....	79
Individual Feature Selection	80
3.1.2. Results.....	82
Significant Textures by Mann-Whitney U-test.....	82
Most Discriminating Textures	83
3.1.3. Discussions and Conclusion	86
3.2. Classification of Bone Marrow Lesion in Knee MRI by Texture Analysis	89
3.2.1. Methods.....	91
Feature Selection for Classification	91
Classification of Slices and Subjects	92
3.2.2. Results.....	95
Feature Selection.....	95
Classification of Slices.....	96
Classification of Subjects.....	98
3.2.3. Discussions and Conclusion	99
Chapter Summary	105

Chapter 4

Advancements for Active Contour-based Automated Bone Segmentation.....	106
Introduction.....	106
4.1. Developing Termination Criterion for Active Contour	107
4.1.1. Methods.....	109
GVF Snake with Contour Resampling and Unit Normal Force Selection	111
Pixel Movement Based Termination Criteria	114
Group Average Difference Termination Criterion	115
Analysis.....	118
4.1.2. Results.....	120
Trend Line Study	120
Time-accuracy Study	124
Evaluation of Termination Methods on Medical Images	129
4.1.3. Discussions and Conclusion	135

Segmentation Time	135
Contour Resampling	136
Number of Iterations	137
Image Size.....	137
A Similar Termination Metric	138
Limitations of Study	138
4.2. Developing Parametric Active Contour with Centroid Force.....	140
4.2.1. Methods.....	143
Image Processing Procedures	143
Images and Analysis	148
4.2.2. Results.....	148
4.2.3. Discussions and Conclusion	152
Chapter Summary	155
 Chapter 5	
Summary and Significance of This Thesis	157
References.....	163

List of Figures and Tables

Chapter 1

Fig. 1.1. Age-adjusted prevalence of doctor-diagnosed arthritis among adults, by sex and body mass index (BMI) category.....	16
Fig. 1.2. Illustration of histological layers of cartilage at the articulating surface of bone	24
Fig. 1.3. Factors contributing to diagnostic decisions.	30
Fig. 1.4. Components of a CAD system	32

Chapter 2

Fig. 2.1. Examples of fast spin echo PD images.	48
Fig. 2.2. An example of the result of cartilage segmentation.	51
Fig. 2.3. An illustration of how the cartilage in one image slice is divided into different sections.....	53
Fig. 2.4. Results of window leveling.	55
Fig. 2.5. Cumulative plot of normal part of the cartilage (in green) and damaged part of the cartilage (in red) derived from the 9 datasets.....	57
Fig. 2.6. 3D plots of the cartilage using different values of γ	58
Fig. 2.7. A cartilage 3D plot together with two MR images	59
Fig. 2.8. Relationship between the difference in Mean and Median and the percentage damaged cartilage presence.	68
Table 2.1. Characteristics of the data used.....	67

Chapter 3

Fig. 3.1. Representative examples of MR knee images and segmentations.	78
Fig. 3.2. Histograms of four most discriminative texture parameters selected using Fisher coefficient and Mahalanobis distance	84
Fig. 3.3. Histograms of four most discriminative texture parameters selected using Euclidean distance	85
Fig. 3.4. Histograms of four most discriminative texture parameters selected using the performance of 1-nearest neighbor classification	86
Fig. 3.5. ROC curves of classifications of Normal and Affected slices using various classifiers.....	97
Fig. 3.6. ROC curves for classification of subjects	99
Fig. 3.7. AUC using NMSC with different number of features	102
Fig. 3.8. Examples of erroneously identified slices.....	104
Table 3.1. Number of parameters from different textural types having statistically significant difference between affected image slice and normal image slice analyzed with Mann-Whitney U-test	82
Table 3.2. Ten most discriminating textures selected by different methods.....	83

Table 3.3. Ten features selected by forward feature selection based on inter-intra class distance	96
Table 3.4. AUC and <i>p</i> -values of the AUC using different classifiers	97
Table 3.5. Classification of subjects using different number of slices as threshold	98

Chapter 4

Fig. 4.1. Illustration of exemplar calculation of GAD.....	116
Fig. 4.2. Comparison of segmentation with pixel movement based evolution with evolution with contour resampling.	121
Fig. 4.3. Trend lines of pixel movement based metrics.....	123
Fig. 4.4. Trend lines of metrics related to Group Average Difference.....	124
Fig. 4.5. Comparison of segmentation results using GAD and APM.....	126
Fig. 4.6. Results of 3 representative segmentations of anatomical structures.	133
Fig. 4.7. Comparison of trend lines using APM and GAD.....	134
Fig. 4.8. Illustration of image gradient and GVF field.	145
Fig. 4.9. Centroid force.....	147
Fig. 4.10. Combination of centroid force and GVF force.	150
Fig. 4.11. Segmentations with and without centroid force.	151

Table 4.1. Comparison of time taken, number of iterations to termination and accuracy for 6 synthetic test shapes using the conventional method and the proposed method....	127
Table 4.2. Comparison of time taken, number of iterations to termination and accuracy for 6 synthetic test shapes using the conventional method and the proposed method....	128
Table 4.3. Comparison of time taken, number of iterations to termination and accuracy for 9 medical images of three anatomical locations using the conventional method and the proposed method in automatic implementation.....	131

Chapter 5

Fig. 5.1. A complete schema for a fully automatic CAD system.	160
---	-----

Statement of Originality

This section is to declare that this thesis is based on the author's original work and all works from other authors were properly acknowledged and cited. The work in this thesis has resulted in original publications including journal article and conference proceedings listed here.

Published journal paper

1. T.K. Chuah, J.H. Lim, and C.L. Poh. *Group Average Difference: A Termination Criterion for Active Contour*. Journal of Digital Imaging, **2012**: **25**(2) p. 279-93.

Published conference proceedings (peer reviewed 4-page papers)

2. T.K. Chuah, C.L. Poh, and K. Sheah. *Quantitative texture analysis of MRI images for detection of cartilage-related bone marrow edema*. 2011. IEEE EMBC, Boston, MA.
3. T.K. Chuah, C.W. Lim, C.L. Poh, and K. Sheah. *Bone segmentation of magnetic resonance images by Gradient Vector Flow active contour with atlas based centroid forces*. 2010. DICTA, Sydney, NSW.
4. C.L. Poh, T.K. Chuah, and K. Sheah. *Differentiating healthy cartilage and damaged cartilage using magnetic resonance images in a quantitative manner*. 2010. DICTA, Sydney, NSW.

Publication accepted

- T.K. Chuah, C.L. Poh, E. Van Reeth and K. Sheah, *Texture analysis and classification of bone marrow lesion in knee MRI - data from the Osteoarthritis Initiative*. Magnetic Resonance Imaging. In press.

Chapter 1

Introduction and Literature Review

1.1. Motivation

Musculoskeletal disease includes any kind of diseases affecting the bone, joint, muscles, tendon and ligament. Osteoarthritis (OA) is a musculoskeletal disease characterized by progressive cartilage wear at the initial stages and cartilage loss at the later stages. Patients of OA suffer from joint pain and stiffness typically at the knee, hip, hand, foot and spine and less commonly at the shoulder, wrist, ankle and elbow [1]. Other disease complications at late and intermediate stages of OA include thickening of subarticular bone, formation of osteophytes, laxity of ligaments, weakening of periarticular muscles, swelling of the synovium and inflammation [2]. The result of OA is the failure of the joint as an organ with permanent disability of the joint.

OA, the most common type of arthritis, is responsible for a large number of disabilities among the world's population [3-6]. It contributes to half of the disabilities in the older population in the US [7]. Among them, women have higher prevalence than men in becoming disabled by OA. Large societal problems can be anticipated with the expected

increase in the prevalence of OA within coming decades [8]. Fig. 1.1 shows the prevalence of arthritis (in the US) with respect to different body mass index [9].

Fig. 1.1. Age-adjusted prevalence of doctor-diagnosed arthritis among adults, by sex and body mass index (BMI) category*† [10].

* Age adjusted to the 2000 U.S. projected adult population, using three age groups: 18--44 years, 45--64 years, and ≥ 65 years.

† BMI = weight (kg) / height (m²). Categorized as follows: underweight/normal weight (<25.0), overweight (25.0 to <30.0), and obese (≥ 30.0).

In Singapore, OA was listed fifth in the top ten causes of disability burden in 2004. The estimated years of life lost due to OA is 4.9% [11]. Loss of healthy life years can be quantified by the loss of disability adjusted life years (DALY). As OA become more prevalent with age and with the aging of current population, loss of DALY due to OA is expected to increase by 40% by 2030 (total DALY lost in 2007 was 17000) [12]. These data raise the need to improve the efficiency of diagnosis and treatments of OA to decrease the societal burden caused by the disease.

Patients affected by OA typically only seek medical attention when joint pain is present. The clinical gold standard for first decision making is based on projection X-ray radiography. This method of diagnosis is fast, inexpensive and non-invasive: once a radiograph is obtained, physician uses the extent of joint space narrowing and presence of osteophyte to tell the seriousness of cartilage degeneration. However, this technique has been known to be insensitive in detecting early degeneration of cartilage [13]. It is also insensitive to changes in bone marrow. Projection radiography gives a high false negative in OA because the target structure (cartilage) affected is invisible and diagnosis decision is based on joint space width of the patient. While projection radiography is able to give positive diagnosis of OA, the disease often has progressed to its irreversible stage and the only treatments are pain management drug or joint replacement surgery.

Diagnosis of OA involves examining multiple structures of the joint. Advancement in magnetic resonance imaging (MRI) enables both the hard and soft tissues of the whole joint to be visualized in great details. To locate cartilage lesion, radiologist typically examines a number of slices in a routine MRI scan. Furthermore, the radiologist also needs to look for other indications of joint disorder such as synovitis, change in trabecular structure, bone marrow lesion, osteophytes, meniscal tear or extrusion and local cartilage thinning without lesion-like appearance. All these current assessment to the knee for OA are mostly qualitative. The shortcoming of using only qualitative evaluation is that the decision could be subjective and variation of diagnosis between different radiologists could be high. New conceptualization of OA in recent years suggests that the focus on cartilage alone in OA diagnosis is insufficient and the knee

needs to be assessed in totality [14]. To understand the whole disease progression, loss of hyaline cartilage is only the most apparent factor in the diagnosis of OA, which is far from the complete picture. Bone remodeling and attrition, degeneration of fibrocartilage, appearance of chondro-osteophytes, lining cell hyperplasia or joint pain may have occurred in conjunction to or even before hyaline cartilage degeneration. It has been shown that combination of biochemical markers and imaging markers can improve diagnosis of OA [15]. To look for and to integrate this large amount of information presents a great challenge to radiologists, particularly when presented with many patients. Moreover, the size of data and the types of sequences have been increasing over the years. Therefore, a computer-aided diagnosis (CAD) system that can support this diagnosis process would be of great value to improving the reproducibility and accuracy of diagnosis. It would provide efficiency and sustainability to the process.

A CAD system is capable of providing second opinion to the radiologist. Although CAD system is not expected to provide the final decision [16], it can introduce quantitative measurements and statistical knowledge with artificial intelligence into the system which could improve the consistency of diagnosis, both for one radiologist or between different radiologists. CAD can also be developed to integrate information from different parts of a 3D image, and provide better visualization of the potentially pathological part of an image. All these benefits could translate into more effective diagnosis and better patient care: early detection of pathology; accurate screening of healthy subjects; and fewer unnecessary tests resulted from false positives. The overall medical cost of the society can thus be lowered.

Motivated by the need to diagnose OA accurately before it becoming irreversible, and the need to handle large amount of complex data efficiently, this thesis addresses research gaps in developing a CAD system based on MRI for diagnosing OA. Although CAD system is fairly established and highly automated in the field of breast cancer detection [17], it is still in research stage for OA. The reason may be that OA affects many joints, and that medical images of the joints are relatively more complex. However, the joint that causes most number of disabilities is the knee. Therefore this thesis will focus on the diagnosis of OA using images of the knee.

1.2. Diagnosis of Osteoarthritis

1.2.1. Symptoms and Biological Changes

The indication of OA can occur at three different levels, namely the *system level*, the *tissue level* and the *cellular and molecular level*. System level indication is defined as the response made by the body as a whole, including body ache and molecular events not local to the site affected. Tissue level indication is defined as the change of bulk structure locally at the site affected. Cellular and molecular indications are defined as the local changes at the affected site that involves either cellular activity or molecular reactions or both. The following paragraphs discuss the progression of OA at these individual levels.

System Level

Conventional diagnoses are carried out based on the fact that the patient having OA has joint pain (typically the knee, hand or hip is affected), morning joint stiffness, crepitus, visible spurs on X-ray and synovial fluid has OA indication (clear, viscous and/or white blood cell count $< 2000/\text{mm}^3$) [18]. All these criteria examine the presence of OA at the **system** level.

Bone spurs, also denoted as osteophytes, are the bony protrusion overgrown to regions of the cartilage or around the cartilage. These protrusions are useful indication of cartilage degeneration. A classification system (the Kellgren-Lawrence grading system) based on presence of osteophyte and joint space width has long been developed and it has been in use for radiographic assessment [2]. Generally, all cases of OA can be grouped into two major types: primary OA and secondary OA. Primary OA is also called idiopathic OA. This type includes all patients with no traceable medical history that can relate to the occurrence of OA. Hence, it is widely believed that occurrence of this type is related to but not caused by aging. Secondary OA are those with known events or diseases that are associated with OA [18]. Primary OA is more difficult to diagnose because it is chronic and appears to be independent of easily observable factors such as trauma or calcium deposition disease.

Tissue Level

Although some research has concluded that systemic response occurs once cartilage degeneration is initiated [19], OA can occur subtly at the **tissue** level (i.e. the cartilage

itself) much earlier than systemic indication become apparent. At the tissue level, the onset of OA occurs when there is fibrillation at the tangential layer of the cartilage. Split lines starts to form parallel to the articulating surface with their orientation similar to the collagen fibres. This *initiation of microcracks* is caused by abnormal kinematics of the joint. Under the loading of the individual, the microcracks will soon propagate and deepen. This process of deepening extends into the radial layer of the cartilage, causing formation of vertical clefts with clusters of chondrocytes surrounding the clefts. This process is called *pitting*. As the disease progresses, pitting and microcracks formation produces bony and cartilaginous fragments which accelerate the abrasive wear of the cartilage [3].

Cellular and Molecular Level

At the **cellular and molecular level**, the major events happening in OA are apoptosis and matrix degradation [20]. Both events occur dependently. Abnormal kinematics and loading of the joint induce internal stress imbalance within the cartilage. This imbalance of stress leads to transient proliferation of chondrocytes, concurrently with the increase of water and proteoglycan content of the extracellular matrix (ECM) [21]. However, the process is an unsuccessful attempt of cartilage repair and is followed by subsequent apoptosis. Apoptosis of chondrocytes inside the avascular ECM is believed to form matrix vesicles that calcify pathologically because there is no means of removing the vesicles [22]. In the microstructural level, the calcification can be deemed as the initiation of cartilage abrasion and osteophyte formations.

It is important to note the differences between osteoarthritis and rheumatoid arthritis. Inflammation and synovial hyperplasia that invades into the bone and the cartilage are the characteristics of Rheumatoid arthritis (RA). Whereas in OA, there is no synovial hyperplasia, the most prominent process is cartilage degradation. The hallmark of RA is the systemic response which is absent in OA. Studies have demonstrated higher levels of some proteins in RA than in OA, these include IL-1 β , IL-6, IL-8, and TNF- α . MMP-3 level in RA also showed up a tendency to be higher than OA [23].

1.2.2. Diagnosis of Early OA using Magnetic Resonance Imaging

MRI is a noninvasive imaging modality that does not have ionizing radiation. It is used to image structures such as brain, chest, abdomen, knee and even whole body. MRI obtains images by applying a strong magnetic field which aligns the hydrogen nuclei across the patient, and then radio frequency pulses are applied to excite the magnetization of the nuclei, in response the magnetic field generated by the nuclei (during relaxation back to the bulk magnetization) produces signals that the receiver of the scanner can detect to construct an image. Magnetic field gradients are used to cause the nuclei at different locations to rotate at different frequencies and phases so that a 2D image or 3D volume can be constructed [24].

Most OA detected from plain radiograph are difficult to treat, they are either treated by symptom control, or surgical replacement the knee. It is thus important to detect OA earlier than the untreatable stage. Early detection of OA involves detecting the degradation process before or during the onset of microcracks formation, when the

cartilage can still be regenerated. MRI serves as an alternative to visualize the changes of more structures of the joint, and it provides a way to detect cartilage thinning and other signs of OA, before joint space narrowing is obvious. MRI enables early detection in two ways: morphological description and quantitative parameters. Morphological description shows the shape of the cartilage, its accuracy of diagnosis is limited by the pixel resolution of the image acquired. For visualization of cartilage, MRI suffers from the laminar effect (i.e. magic angle effect). It also suffers inability to clearly delineate both the tidemark (Fig. 1.2) and the border to the joint space simultaneously in a type of sequences [25]. Laminar effect is the undesired layered appearance of cartilage due to the variation in the orientation of collagen fiber relative to the applied magnetic field.

Quantitative MRI for provides measures that are related to the composition of the cartilage. In quantitative MRI, the requirement for pixel resolution is not as stringent as that in morphological MRI because morphological data is only used for delineation of region of interest (ROI) so that quantitative parameters can be shown on the correct structure for visualization.

Fig. 1.2. Illustration of histological layers of cartilage at the articulating surface of bone [26]. (reprinted with permission from Springer, license number: 3060700065275).

Morphological MRI

Morphological MR imaging uses primarily fat-suppressed T_1 -weighted spoiled gradient recalled acquisition at steady state (SPGR). Cartilage appears bright in this kind of images. Other candidate sequences include: steady state free precession (SSFP), IDEAL SSFP, fluctuating equilibrium MRI (FEMR), driven equilibrium Fourier transform imaging (DEFT), phase sensitive SSFP, and double echo steady state imaging (DESS) [27-28].

To examine a joint morphologically for early OA (typically knee OA), several parameters are of interest: cartilage thickness at different anatomical regions, total cartilage volume, total cartilage surface area, total subchondral bone area, surface curvature, lesion depth and lesion size. These measurements are purely geometrical. Hence, their accuracies are

limited by the voxel size. The finest spatial resolution of a voxel is generally larger than 100 μ m, which marks the limit of spatial resolution of routine clinical MRI. Moreover, the precisions of geometrical measurements are also limited by the segmentation technique applied. Several groups have reported fully automatic schemes to delineate the cartilage with considerable success [29-31]. The problem of repeatability of segmentation does not seem to be a bottle neck in this field. The challenge remaining is which methods of segmentation can be best applied in a diagnosis system for most efficient and accurate diagnosis decisions.

Some advanced techniques in morphological MRI have been proposed. Qazi suggested that diseased cartilage is more homogeneous and homogeneity is able to separate healthy and diseased cartilages [32-33]. Tamez-Peña used statistical parametric mapping to visualize the change of cartilage thickness in follow-up study [34].

Quantitative MRI of the Cartilage Composition

Reducing the voxel size of a magnetic resonance (MR) image involves the reducing contrast-to-noise ratio or signal-to-noise ratio and increasing acquisition time. Therefore the voxel size cannot be reduced indefinitely in practice. It was recommended that for a 3T imaging system for assessment of cartilage, a good balance between image resolution and acquisition time can be obtained with 1.5 mm slice thickness with 0.3 mm isotropic in-plane resolution [35]. Although the image resolution is limited, molecular information of the cartilage such as collagen content, proteoglycan (PG) content and abundance of water can be estimated with quantitative MRI for cartilage composition. This mode of

imaging does not require a very precise delineation of the cartilage yet it is able to display variation of molecular architecture. Important quantitative MRI sequences that image composition of cartilage include T_2 , $T1_\rho$ and $T1_{Gd}$ [27]. Sodium MRI is also able to image for glycosaminoglycan (GAG) concentration.

T_2 map is an image that shows the transverse relaxation time of tissue at each pixel position. Tissue hydration affects T_2 value (on T_2 maps) significantly. Increased T_2 value was found not only on areas of cartilage degeneration but also linked to decrease of both cartilage volume and thickness [36-37]. Changes in collagen content have also been shown to affect T_2 value. The sensitivity of T_2 to tissue hydration made T_2 a potential tool in detecting inflammatory changes because of the increased osmotic pressure associated with cartilage. Changes in water content of patellar cartilage in healthy subjects due to exercise was detected using T_2 maps [38]. Moreover, increased mobility of water in degenerating cartilage produces increased T_2 value. One shortcoming of T_2 images is that currently it is unable to tell clearly between aged cartilage and OA cartilage.

$T1_\rho$ has been known to detect the disruption of integrity of cartilage matrix because increase in the measured $T1_\rho$ is associated with the loss of proteoglycan [39-40]. Proteoglycan contributes to the fixed charge density (FCD) of the cartilage due to its negatively charged side groups, the GAG. Loss of proteoglycan alters the FCD. It was demonstrated that the value of $T1_\rho$ correlated well with FCD [41]. Moreover, collagenous component of the cartilage is often intact prior and during the onset of OA while loss of

proteoglycan has already started. This information suggests that $T1_{\rho}$ gives earlier indication than techniques to probe for change in collagenous components. $T1_{\rho}$ is deemed to be more accurate than T_2 in discriminating healthy and OA-affected cartilage, also due to its higher dynamic range [40].

$T1_{Gd}$ is able to give a quantification of proteoglycan loss because the fixed charge density of cartilage is decreased in diseased cartilage so diseased cartilage absorbed more Gd-DTPA²⁻ [42-43]. This imaging technique is called delayed-Gadolinium-Enhanced MRI of the Cartilage (dGEMRIC). This technique has demonstrated low inter and intra observer variability, suggesting its feasibility in clinical practice [44]. However, the study by Hori showed a wide scatter of blood levels both across subjects and within subjects with similar body mass index (BMI) [45]. This wide scatter was believed to cause over-pronounced severity of cartilage damage in some patients. Therefore, the accuracy of the contrast-dependent technique bears the disadvantage that the image is dependent on the dose of contrast administered between different subjects and patient's tolerance to the contrast agent. It is unclear whether the contrast uptake of similarly healthy subjects is the same given that their blood levels of contrast agent vary with time at the same rate. Moreover, Taylor found there is no correlation between precontrast $T1_{\rho}$ and $T1_{Gd}$ for excised cartilage specimens [40]. These findings complicate the notion of association of a sequence to a particular composition. Further investigation is necessary to clarify the complication. Imaging consistency and complication can be reduced by imaging without a contrast agent.

MRI of Other Tissues

The source of pain in hip OA has been investigated by studying MRI of the brain [46]. Meniscus and ligament draw clinical attention primarily when there is sign of tear. Morphological MRI is mainly used to study meniscus and ligaments. Compositional imaging using quantitative MRI, which is widely studied for cartilage, is rarely used for the diagnosis of meniscus and ligament possibly because they appear relatively small compared to other structures and the microstructures (such as fibrocartilage) in them align in a complex manner. Furthermore, intact meniscus appears mostly homogeneous in most MR images. Other than the articular cartilage, compositional imaging is useful mostly for visualizing changes in molecular level for the bone marrow.

Bone marrow edema (a.k.a. bone marrow lesion, in short BML), a main focus of this thesis, is indicated by focally increased signal in water-weighted sequences including T₂-weighted fat saturated sequences, and proton density (intermediate-weighted) fat saturated sequences. MRI is an important imaging modality that can be used to visualize BML without the use of contrast agent, which is currently not possible by other modalities (including X-ray, computed tomography and ultrasound). When bone marrow fat is replaced by material containing H⁺ ions, BML that carries the signal of water can be imaged by MRI [47]. MRI produces volumetric dataset which allows multi-planar analysis of the BML. BML can be assessed either quantitatively [48], or semi-quantitatively using scoring system [49].

BML appears in many diseases of the knee. These include osteoarthritis (OA), arthritis, complex regional pain syndrome, osteonecrosis, bone bruise, osteochondritis dissecans, and even tumor or transient bone marrow edema syndrome [50]. It was reported that knee pain is related to increasing size of BML [51], and the size of BML has been associated with the risk of cartilage loss [52-53]. Furthermore, BML has been found to be a strong risk factor for structural deterioration in knee osteoarthritis [54]. Such a common indication of musculoskeletal disease is worth studying and developing into computer-aided diagnosis system. More details of BML are presented in Chapter 2.

1.3. Overview of a Computer-aided Diagnosis System for musculoskeletal diseases

A computer-aided diagnosis (CAD) system is designed to provide complementary second opinion to the radiologist or clinician so that the diagnostic decision could be made more accurate and objective. It is intended to decrease the possibility of personal bias across different clinicians/radiologists. The design of the system does not intend to take over the task of final decision making in diagnosis [16]. It is important to note the main difference between CAD system and automated computer diagnosis system: the former (CAD) does not need to be better than or comparable to the ability of a radiologists but only needs to provide information capable of improving radiologists' efficiency or accuracy of decision; the latter is intended to take over radiologists' tasks and its use would not be justified unless its performance is better than or comparable to that of radiologists.

As shown in Fig. 1.3, the image based CAD system could aid visualizing subtle changes, provides easy and fast reading of image information and improving consistency of radiologists' report. Final diagnostic decision usually depends on both the radiologist's report and other data such as medical history and physical data. Medical history such as bone trauma and cases of cancer in family members can make a difference in the diagnosis decision. Physical data such as height, weight, race and existence of pain also contribute to important consideration of diagnosis decision.

Fig. 1.3. Factors contributing to diagnostic decisions.

Fig. 1.4 shows the components in a typical CAD system for musculoskeletal diseases. It consists of processes such as image processing, feature analysis and data classification [55]. Segmentation of structures such as the bone and the cartilage, is a challenging and indispensable part of the CAD system. It provides information to be extracted for analysis. This information is basic but not trivial to obtain. Manual segmentation process is time consuming thus automatic or interactive segmentation processes have been actively developed in the field. Once the segmentation is obtained for each structure, analysis of the structure can be carried out either by the computer or by the radiologist. To enable

radiologist to analyze information quickly, appropriate 3D visualization system has to be developed. It should enable the radiologist to switch between 3D information and the conventional slice-by-slice view of the image. The 3D information that can be incorporated into the visualization system includes cartilage thickness map which highlights potentially pathological areas. Parallel with this visualization system is the classification system which deduces the state of disease of an instance. It either gives the probability of being affected by certain disorder or discretely classifies the instances into different disease stages. This result may aid the radiologists in 2 ways: (i) to screen healthy instances and (ii) to make an objective and informed decision for ambiguous instances.

An image-based analysis system analyzes the features from images. Those features include intensity, texture, and geometrical measurement such as cartilage volume, cartilage thickness, and size of bone marrow lesion. With prior knowledge of these features in a population, classification models can thus be developed to provide statistical knowledge and objectivity into the CAD system. In addition to image-based CAD, other data obtained from the patient can also be incorporated into the computerized system to calculate the probability of having certain disorder.

Fig. 1.4. Components of a CAD system.

1.4. Image Understanding of Knee MRI

Image understanding or machine vision of medical image is the artificial ability of computers to translate image information into clinically relevant information, such as state of disease or possibility of disease. It is a kind of artificial intelligence and the outcome is obtained based on statistical data such as past instances. It has the potential of correcting or complementing diagnostic decisions that are prone to judgment or personal

bias. This section (1.4) reviews and discusses studies related to the development of image understanding of knee MRI.

1.4.1. Medical Image Segmentation and Registration

As shown in Fig. 1.4, to obtain quantitative measurements, such as cartilage thickness or cartilage volume, from medical images, segmentation of structures such as bone, cartilage, muscles, ligaments and menisci, is necessary. Segmentation is the process of extracting target structures and removing undesired structures including the background. Computerized segmentation methods are known to be able to handle large number of data, and also have been validated to be more reproducible than manual segmentation. An example would be the segmentation of cartilage using B-spline snake which showed better reproducibility [56]. Different types of segmentation methods have been developed for medical images, they include thresholding, region growing, classifier, clustering, deformable models and atlas guided approach [57]. A review of segmentation for the knee was published by Sun et al. [58] in 2006. Registration of images is employed mostly to align different acquisitions of the same patient, or for aligning data from different patients for comparison. Surveys of image registration methods were published by Pluim and Zitová's groups [59-60]. Segmentation and registration are processes closely related to each other: the information resulting from registration can be used for segmentation; and the information resulting from segmentation can be used for registration. Simultaneous segmentation and registration algorithms have been proposed and developed but their use is still limited to research [61-63]. However, to the best of the author's current knowledge, this kind of algorithm has not been reported for segmenting structures in knee MRI.

Since the cartilage is the part that is worn out and needs considerable attention in diagnosis of OA, much work has been done to segment the cartilage using knee MRI. Research has been ongoing in the development of semi-automatic and automatic segmentation methods. Ghosh et al. used watershed algorithm to segment cartilage from high resolution MRI [64]. Dam et al. developed a semi-automatic method that involves voxel classification, with an interactive step that combines the posterior probability map from classification and watershed transformation [65]. Fripp and Part reported an automatic segmentation algorithm for cartilage of the knee: the algorithm uses active shape model (to extract the bone), likelihood estimation of bone cartilage interface, and cartilage thickness model [66-67]. These works reported encouraging results but they studied only the segmentation using T_1 -weighted (or SPGR) sequences of knee MRI. It is important to point out that different sequences have different advantages in visualizing certain structure, but also different challenges in developing segmentation algorithms. Another sequence that has been studied much in cartilage segmentation is DESS (Dual Echo Steady State) because it is able to enhance the contrast between cartilage and synovial fluid. Shim et al. has developed a semi-automatic segmentation method based on graph cuts [68] and Dodin et al. reported an automatic segmentation method that involves texture analysis and Bayesian decision criterion [69].

Most studies that developed segmentation of the structures at the joint still only analyze the cartilage. However, OA could have started at the molecular level (which may be observed from the bone marrow using T_2 images) even before cartilage thinning is

observable. To increase the accuracy of diagnosis, visualizing the bone and bone marrow is also necessary [14]. A commonly used sequence to visualize multiple structures is proton density (PD) or intermediate-weighted (IW) sequence [70]. They provide better visualization of the bone marrow as compared to the sequences that are optimized for cartilage imaging. Therefore, to develop a CAD system that can integrate information from different structures, it is necessary to segment bones from PD, IW and also T₂-weighted images that provides good visualization of bone marrow lesion.

One approach to segment bones and other structures from those sequences would be to utilize the automatic segmentation already available to other sequences. It is known that automatic segmentation of the cartilage from DESS [69], and automatic segmentation of the bone and cartilage from fat suppressed SPGR [66, 71] have been reported with considerable success. It then seems straightforward to register the segmentations obtained from these sequences for segmentation of structures in the sequences of interest (i.e. PD, IW and also T₂). However, directly applying registration has several problems:

- The sequences that can be segmented automatically may not contain enough information to characterize certain target structure, for example, bone marrow lesion or osteophytes. Hence, direct registration may not work in a straightforward manner.
- Because the knee consists of both rigid (bone) and deformable (cartilage, meniscus, ligament, synovium, fat, muscle, skin) structures, neither rigid nor deformable registration alone can provide accurate registration result.

Due to the problems mentioned, registration across modalities or across different sequences of MRI possesses its own challenges. When the task on hand is to analyze certain structure based on certain MR sequence, it may be more feasible to segment target structures using those different sequences directly instead of attempting to find the best registration method to segment the target structure from another sequence.

Component Algorithms for MRI Knee Segmentation

A non-trivial segmentation task usually involves more than one segmentation algorithm to complete, for example, initialization can involve generalized Hough transform or thresholding, and refinement of final segmentation can involve constrained active contour. Here we denote the algorithms involved in a segmentation task as component algorithms. Developing an automatic or semi-automatic segmentation for a particular sequence involves choosing component algorithms, developing the algorithms, and testing and optimization of parameters in the component algorithms. Statistical Shape Model (SSM) is typically used as a main component algorithm for segmentation of knee structures [71-73]. The use of SSM has been shown to be effective [74] for sequences similar to SPGR. Its advantage is that the shape of final segmentation of a certain structure would not deviate too much from the expected shape. However, the generation of the SSM itself is a demanding task requiring resources such as: many instances of 3D knee MRI, high memory fast computational system, and possibly long computational time. Furthermore, the model creation and the use of the model is an extensive research topic. The resource requirements for implementing SSM based method is a potential hurdle to practicality.

An alternative component algorithm pursued in this thesis is the use of active contour that derived its constraints from a single reference subject. It is believed that with proper constraints and termination, active contour can be applied for automatic segmentation of the knee structures. The next subsection (1.4.1.2) reviews the subject of active contour and Chapter 3 presents advancements on constraining the active contour for knee bone segmentation, as well as terminating the contour with better efficiency.

Active Contour Segmentation

Active contours or snakes [75] are moving curves used for detecting image boundaries. It is mostly applied in segmentation tasks: these include segmentations applied on radar imagery [76], computed tomography [77], magnetic resonance images of the muscles [78] and ultrasonic breast images [79]. Active contour can be broadly categorized into two types: parametric and geodesic. Parametric active contour employs an energy functional minimization formulation to search for the correct location for a parameterized contour [75]. Hence, it is more suitable for segmentation of single structure in an image. Geodesic active contour uses level set methods to find the locations of final contours. It is more suitable for segmentation of multiple structures in an image due to its robustness in merging and splitting the contours during evolution [80].

The earliest active contour was parametric. It depends on intensity gradient on the image and it has fixed number of sampling points. The disadvantages result from these are that the contour does not move if it is initialized too far from the actual object boundary, and the contour cannot have too much concavity nor too much size difference from the

initialized contour. Solutions to these problems were then developed. To enable the contour to move even when it is initialized far away from the actual object boundary, Cohen proposed the use of balloon forces [81]; Xu and Prince proposed the use of gradient vector flow [82]. To enable more robust size and concavity change, Lobregt and Viergever suggested a way to resample the contour during its evolution step. Merging [83] and splitting [84] of parametric active contour have also been developed and tested.

The use of topological parametric active contour has the advantage that it is easy to formulate the energy minimization framework [84]. However, implementation of geodesic active contour based on level set is equally robust and has been widely used, perhaps due to its simplicity of implementation. Level set active contour offers a robust alternative to merging and splitting active contour because the evolving function is one dimension higher than the contour itself, making merging and splitting intuitive, without need of setting additional conditions. Chan and Vese developed active contour based on level set, using region information that does not need the presence of sharp edge [80]. Li et al. proposed a method for removing the re-initialization step in evolving a level set active contour [85]. A fast level set based method without edges and without re-initialization was demonstrated and tested by Zhou et al. [86]. Zhou's method is a fast and jumpy version of active contour: it does not evolve slowly and smoothly like Chan's or Li's snake, its advantage is speed but it also carries the risk of higher error for some applications.

For completeness, the formulations for Chan's and Li's active contours are shown as follow.

$$\textbf{Li's Snake: } \frac{\partial \phi}{\partial t} = \mu \left[\Delta \phi - \operatorname{div} \frac{\nabla \phi}{|\nabla \phi|} \right] + \lambda \delta(\phi) \operatorname{div} \left(g \frac{\nabla \phi}{|\nabla \phi|} \right) + v g \delta(\phi) \quad (1-1)$$

$$\textbf{Chan's Snake: } \frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mu \operatorname{div} \frac{\nabla \phi}{|\nabla \phi|} - v - \lambda_1 (u_0 - c_1)^2 + \lambda_2 (u_0 - c_2)^2 \right] \quad (1-2)$$

Where

ϕ is short for $\phi(x, y, t)$ for a contour on 2D image that evolves with time, it is the level set function that defines the contour.

Δ is the Laplacian operator.

div denotes divergence operator, ∇ is the gradient operator, and $\frac{\nabla \phi}{|\nabla \phi|}$ denotes

the normal of the level set function.

$\delta(\phi)$ is the Dirac delta function of the level set function.

g is $g(x, y)$ which is the edge function of the image, it is very small (near to zero)

at edges and positive at homogeneous region.

u_0 is $u_0(x, y)$ which is the grey value in the 2D image.

c_1 is the average gray value inside the contour, and c_2 is the average grey value outside the contour.

Practical segmentation tasks can have more than 2 homogeneous regions (foreground and background) in the same image. Consequently, any conventional 2-phase active contour is challenged. To segment images with more than 2 regions using active contour, multiphase level set method should be employed. Multiphase method was introduced by Vese and Chan [87]. Improved version of multiphase active contour has been tested with knee MRI and CT images [88], however, the study was done in 2D and was preliminary in nature. The focus was to test the speed of active contour instead of applicability to segment multiple slices and multiple volumes. There is much room for development of the method into the regime of CAD.

1.4.2. Feature Analysis of Anatomical Structures

Features of the anatomical structures, such as cartilage thickness [89-90], cartilage texture, cartilage intensity, cartilage homogeneity [91] and size of bone marrow edema [92] are all important for deciding the diagnostic outcome.

Cartilage geometry is a widely studied feature for knee MRI. It has been found out that MRI can provide precise and accurate analysis of the morphometry of cartilage [93-94]. The coefficient of variation of cartilage volume has been found to be low between repeated MRI scans for the same subject [95]. Muensterer et al. have segmented the cartilage semi-automatically and measured geometrical parameters of the cartilage from MRI [96]. They found good correspondence with anatomical section or arthroscopic measurement. Haubner et al. validated the measurements of cartilage thickness with CT arthrography and confirmed that MRI provides comparable measurements to CT [97].

For visualizing the cartilage thickness, computer frameworks have been developed to display cartilage thickness map [98-99]. LiveWire segmentation method was tested in measuring cartilage thickness [100], it has been shown to be more repeatable and faster than manual segmentation, but did not approximate the true thickness as well as manual segmentation. The co-occurrence matrix-based texture of cartilage on T1 ρ map was studied and has been found to be potentially useful in longitudinal and population studies of knee OA [101].

Characterization of bones using MRI by texture analysis of the bone has been studied fairly extensively, especially in the characterization of osteoporosis. Using high resolution MRI, histomorphometric texture measures have been found to be effective in differentiating subjects with and without osteoporotic bone fracture [102-103], estimation of bone strength [104-108], quantification of bone structures [109] and detection of osteoporosis [110]. However, these studies mainly focused on the pattern and structure of the trabecular bone (histomorphometric texture measures) instead of the overall appearance of the bone marrow. There is a range of TA methods for example structural-based, model-based, statistical and transform-based, with statistical methods being the most widely used technique in medical images [111]. A recent study showed that textural features based on statistics, model and wavelets were able to discriminate different tissues such as bone marrow and fat in the knee joint [112].

Using features derived from images, classification of normal patients (i.e. controls) and patients suffering from disease (i.e. cases) can be performed. For example, trabecular

bone texture in MRI and CT was used to distinguish between osteoporosis cases and control, where three features were examined and 90% of the cases were well estimated [113]. Textural features were able to discriminate different tissues including bone marrow [112].

1.5. Specific Objectives

The overall objective of this thesis is to make advancements to the development of the CAD system for detection of musculoskeletal diseases such as OA using knee MRI, focusing on visualization of defective cartilage, analysis of textural measures and classification based on textural features for BML detection, and active contour based method for automated segmentation of the bone.

Chapter 2 aims to investigate potentially useful parameter for detecting cartilage lesion and to develop a visualization system to aid the study of cartilage lesion. Section 2.1 presents the visualization system developed for viewing cartilage and efficiently highlighting areas suspicious of having lesions. Section 2.2 studied the presence of cartilage lesion based on histogram-derived parameters. A good correlation between the parameter and the presence of lesion has been shown and can lead to a successful use as an indicator for presence of cartilage lesion in CAD.

Chapter 3 aims to discover and determine which texture parameters (in water sensitive MR sequences) have potential to be incorporated into the CAD system. Section 3.1

presents a study into the possibility of quantitative analysis for detecting the presence of BML. Here we hypothesized that the presence of BML can be detected and analyzed quantitatively and consistently using texture analysis. By examining the separability of texture parameters for distal femoral marrow with and without BML, we determined which parameters support the hypothesis. Following the discovery of useful textural parameters for analysis of BML, section 3.2 investigates texture based classification of MRI slices with and without bone marrow lesion (BML). This section will demonstrate the feasibility of using texture based analysis to classify slices and subjects.

The objective of Chapter 4 is to present advancements made in the application of active contour based segmentation for the segmentation of knee bones. These studies aim to overcome the hurdles of applying active contour effectively for knee bone segmentation. Section 4.1 presents a study to improve the efficiency of termination of active contour on knee MRI images. The aim is to remove tuning processes and manual determination of termination parameters so that the application of active contour can be more automated. The section also reviews presently used termination criteria. The constraint needed in applying active contour successfully for knee bone segmentation is presented in section 4.2. This section proposed the use of single-subject based centroid force to constrain the movement of the contour at ambiguous boundaries in PD images, and evaluated the accuracy of the method.

Chapter 5 concludes the studies and also relates the chapters to the overall aim. It states the significance of the research presented in this thesis and how it may be helpful to

research scientists, clinicians, radiologists, companies developing CAD system, and to the general public.

Chapter 2

Visualization and Image Analysis of Damaged Cartilage using Knee Joint MRI

Introduction

As the cartilage is the most apparent sign of OA, 3D visualization of the damaged cartilage is important to aid the understanding of cartilage damage or thinning with respect to disease progression. The visualization system could make diagnosis of cartilage damage more efficient. This chapter focuses on the study of cartilage appearance using proton density-weighted images and developing a novel visualization system which is designed to highlight areas suspicious of cartilage lesion. This chapter first presents the development of the visualization framework followed by a study that investigates whether there is a relationship between the signal intensities of the cartilage and the size of the damaged cartilage presence.

2.1. Visualization of Damaged Cartilage of the Knee Joint Using Magnetic Resonance Imaging

Cartilage damage is often caused by injuries or degenerative joint diseases such as OA. OA affects a large proportion of the elderly population and is responsible for a majority of the disability [3-6, 114]. It is important to be able to visualize damaged cartilage during diagnosis using medical imaging to ascertain the size and locations of the defects [115-118]. Location and size of the damaged cartilage provide information for determining the treatments required [119] and for tracking the progression of the defects. MRI is increasingly being used clinically to assess damaged cartilage because it can provide high resolution three dimensional (3D) volumetric images with excellent soft tissue contrast [42, 98, 120-124]. Using MRI, it is possible to visualize and study the cartilage in 3D.

A number of different MR sequences are currently being used clinically in the examination of cartilage damage. These sequences include fat suppressed Spoiled Gradient Recall (SPGR), T_2 -weighted, and fast spin echo proton density (PD) weighted. For example, cartilage can be clearly visualized because it appears bright while the rest of the surrounding tissues (e.g. bone, synovial fluid) appear dark in fat suppressed SPGR MR images [98]. Using this type of images, a number of automated image processing techniques have been developed to derive important parameters of the cartilage (e.g. signal intensities, thickness, volume and surface area) and to visualize the cartilage using 2D/3D thickness maps [95-98, 121, 125-129]. However, because of a full evaluation of the knee joint including subchondral bone, menisci, ligaments as well as surrounding soft

tissues [14], PD weighted images which provide better contrast between the various anatomical structures (including contrasts between fluid, subchondral bone, and cartilage) are generally used in routine protocols [130-132]. In PD images, the cartilage appears grayish as compared to the bone marrow which appears bright (see Fig. 2.1). Healthy cartilage has more homogenous pixel intensities whereas damaged cartilage can have abnormally bright pixel intensities. The abnormally bright pixels generally indicate that there is high fluid content in the cartilage (for the sequence parameters used in our study). This became evident when research found that interstitial water content in cartilage specimens agree with appearances of PD acquired from gradient echo [133].

In current clinical practice, radiologist generally examines the cartilage by visual inspection of the MR dataset in a slice by slice manner and report the damage qualitatively. The radiologist will need to mentally map the locations of the cartilage defects if presence. Therefore, a tool that allows radiologists a quantitative visualization of damaged cartilage using PD sequences and a 3D model would be of great help.

This section (2.1) presents a novel methodology to interactively visualize the damaged portions of the cartilage using clinical PD weighted MR knee images. The study first investigated the pixel intensity distributions of normal femoral cartilage and femoral cartilage with focal damage using PD weighted images. Using the results obtained from the pixel intensity study, a region-based visualization technique to visualize the damaged part of the cartilage in 3D was developed. A method was developed to automatically derive threshold value that can differentiate the damaged cartilage and healthy cartilage

for visualization. This method derives threshold value using best fit Gaussian curve of the cartilage histogram. The whole cartilage is then plotted in 3D using a calibrated colorbar (based on the threshold value) to visualize damaged part of the cartilage.

Fig. 2.1. Examples of fast spin echo PD images. (a) An image of normal cartilage. (b) An image with focal cartilage damage (circled).

2.1.1. Methods

The methodology involves (i) segmentation of the femoral cartilage using pre-processed MR images, (ii) Gaussian curve fitting of pixel intensities to determine the threshold value required to differentiate normal and damaged cartilage, and (iii) visualization of the damaged portions of the cartilage using a custom built interface. All the methods were

developed using MATLAB software (The Mathworks, Inc., Natick, MA, USA). The detailed steps are described in the following sub sections.

Pre-processing and Segmentation of Cartilage

Clinical fast spin echo MR PD weighted volumetric dataset of patients suffering from focal cartilage damage were analyzed in the study. The images were acquired using a 1.5T MR scanner (General Electric Medical Systems, Milwaukee, WI) using a dedicated knee coil. The images were first processed using an automatic window-leveling algorithm [134]. This is to improve consistency in the pixel intensities among the images, to enhance the contrast of the images, and for pixel intensities comparison between different datasets. Based on histogram of each image, the window value is defined as from zero to the 99th percentile of the histogram and the value of the level is defined as the mid-point of the window. This window/level provides good contrast between the cartilage and other tissues. Using the window and level, the pixel intensities of each image are mapped onto an 8-bit gray scale.

After applying auto window-leveling to all the images within a dataset, the femoral cartilage was manually segmented using a custom built software, used in our previous study [98]. This segmentation process involves the users clicking a number of points along the boundary of the cartilage. The points are then connected using cubic spline interpolation. The cartilage segmented in this step includes both the damaged and normal portions, which is subsequently used in the visualization interface. The main advantage of including both portions is that it allows the volume of damaged/lost cartilage to be

calculated. This segmentation process requires prior knowledge in determining how the damaged part of cartilage should appear when it was normal. Because the focus is on focal cartilage damage, an experienced radiologist relies on experience to determine the area of damage.

To analyze the differences in pixel intensities between the normal and damaged portion of the cartilage and to evaluate the accuracy of the visualization technique, the damaged part of the cartilage was also separately segmented. To obtain the segmented image of the healthy portion of the cartilage, the mask of the damaged cartilage is subtracted from the whole cartilage segmented. As a result, there are three types of segmented images – (i) whole cartilage that comprises both healthy and damaged (CW), (ii) damaged portion of the cartilage (CD) and (iii) healthy portion of the cartilage (CH). An example of the cartilage segmentation is shown in Fig. 2.2. The figure shows the masks created for the normal part of the cartilage and the damaged part of the cartilage.

Determination of Threshold Value using Gaussian Curve Fitting

It was observed that the histogram of the pixel intensities of the CW generally follows a Gaussian distribution. Majority of the Gaussian curve consists of the pixel intensities from the normal part of the cartilage because the damaged cartilage tends to be of higher intensities and smaller in numbers for the type of damage studied. A best fit Gaussian curve is used to determine the value of the threshold that can differentiate between the normal and the damaged part of the cartilage. The histogram of the whole cartilage (CW) is fitted with a Gaussian curve, by using an algorithm available at MATLAB Central

[135]. The probability $Pr(Z \leq T) = \gamma$ is then calculated using the equation of the best fit Gaussian curve obtained, where T is the threshold to be derived, γ is the percentage area (e.g. 95th percentile) under the normal distribution curve to the left of T (left-hand tail), and Z is a normal variable. In PD weighted images, damaged cartilage tends to have significantly larger pixel values than normal cartilage due to the intrusion of fluid into the cartilage matrix where the cartilage is damaged [115]. Because these large pixel values generally lie outside the upper limit of the Gaussian curve, majority of the damaged cartilage pixels are expected to be larger than the derived threshold value T . Hence, the damaged cartilage can be differentiated from the healthy cartilage.

Fig. 2.2. An example of the result of cartilage segmentation. The damaged part of the cartilage was included as part of the whole cartilage which will be used in the visualization technique.

Visualization of the Damaged Portion of the Cartilage

A region-based visualization technique was developed for viewing the damaged portion of the cartilage in 3D. The technique was incorporated into an interface to allow clinician to refer back to the original MR images corresponding to points of interest by interactive clicking on the 3D plot. In the first stage of the visualization method, using the segmented images of the whole femoral cartilage (CW), the coordinates of the pixels of the inner cartilage is fitted with a cylinder using a cylinder fitting algorithm [136]. The fitted cylinder is used to estimate the centroid of the femoral cartilage which serves as the center to extend radii onto the cartilage to detect the inner cartilage and the cartilage surface, and to divide the cartilage into regions. As illustrated in Fig. 2.3, the arc angle α between the beginning of the cartilage and the ending of the cartilage is divided equally based on the number of regions λ defined. Using the angle β derived for each region, the cartilage in each image slice is divided into regions. The same number of regions is maintained for the cartilage in all the image slices. This dividing of cartilage in each slice is performed automatically for all the images in a dataset. The average pixel values for the regions are used to plot the 3D cartilage plot. The purpose of the use of regions is to reduce errors that might be caused by the scattered large fluctuation in the pixel intensities caused by noise that will affect the visualization of damaged regions. The color of the 3D plot corresponds to the mean pixel value of each region which represents either normal or damaged portion of the cartilage. The color bar of the plot is calibrated using the threshold value determined by the Gaussian fitting step. Regions with mean pixel intensity smaller than T is given one color (i.e. blue) whereas regions with mean

pixel intensity larger than T is given a shade of different colors. As a result, damaged portion of cartilage will appear in different colors and it will stand out from the normal portion of the cartilage. Using this scheme the damaged cartilage will be highlighted which will enable clinician to clearly visualize the damaged region in 3D. Because there is often a need to refer back to the MR images when the clinician is studying the 3D cartilage plot, a useful feature (3D-Trackback function) was developed. The function enables the user to click on the cartilage 3D plot and the corresponding MR image will be displayed and the corresponding point will be marked on the MR image and the 3D plot for easy reference.

Fig. 2.3. An illustration of how the cartilage in one image slice is divided into different sections. This was performed automatically for all the slices in a segmented volumetric dataset.

Analysis

The difference in the mean pixel values between the damaged and normal cartilage is analyzed using Student t -test, and $p < 0.05$ is considered statistically significant. Accuracy of the thresholding method in differentiating the normal portion of the cartilage and the damaged portion of the cartilage is assessed using the manually segmented images of the damaged and normal cartilage, respectively. In this study, the manually segmented images verified by the experienced radiologist are considered as the ground truth. Percentages of normal and damaged cartilage identified against the ground truth were calculated to assess accuracy.

2.1.2. Results

Nine retrospective knee MR volumetric datasets of patients with focal damage at femoral articular cartilage were studied. The datasets were acquired at sagittal planes with slice thickness ranges from 1.5 to 1.8 mm, with in-plane resolution of either 512×512 or 348×384 pixels. Each of the datasets has focal damage at the femoral cartilage and the degree of cartilage damage ranges from Grade 1 to Grade 3 as determined by an experienced radiologist. The age of patients of the datasets ranges from 37 to 53 years old with mean age of 46 years old. After applying auto window/level, all the MR images were manually segmented and verified by the experienced radiologist. Fig. 2.4 shows an example dataset before and after applying auto window/level. Using auto window/level, the inhomogeneity of pixel intensities between the different slices within a volume dataset (likely to be caused by crosstalk) were made more uniform (see Fig. 2.4). The

histogram obtained after auto window leveling shows that the distribution became more normally distributed. As a result, the differences in the overall intensities from one slice to another within a dataset and the differences between different datasets were significantly reduced. This enabled the different datasets to be better compared.

Fig. 2.4. Results of window leveling. (a) Histogram of all the pixels from one volume dataset before auto-window leveling was applied. (b) Histogram of the pixel intensities after auto window leveling was applied and a more normally distributed curve was obtained. (c) Example images before auto window leveling was applied. (d) Corresponding images from (c) after auto window leveling was applied. A more uniform distribution of pixel intensities can be observed.

In the visualization of the damaged part of the cartilage, the whole cartilage (CW) was used. The damaged part of the cartilage was separately segmented to analyze the differences in the pixel intensities between the normal part of the cartilage and the damaged part of the cartilage, and to assess the accuracy of the thresholding technique in differentiating normal and damaged cartilage. Using the segmented images (CD and CH),

the average volume of the damaged cartilage studied was $505 \pm 174 \text{ mm}^3$ and the percentage of damaged cartilage over whole cartilage was $5.39 \pm 1.69 \%$.

Fig. 2.5 shows the cumulative histogram of the normal part of the cartilage and the damaged part of the cartilage, derived using the 9 samples of cartilage with damage. The histogram of the normal part of the cartilage displays a normal distribution, with pixel intensities generally smaller than that of the damaged part of the cartilage. Results show that there was a significant difference between the mean pixel intensities of the normal part of the cartilage (87.5) and the mean pixel intensities of the damaged part of the cartilage (183.4) ($p < 0.001$). This result suggests that it is possible to differentiate between the normal and the damaged part of the cartilage using a threshold value.

The segmented images of the whole cartilage (CW) were used to derive the threshold value required to differentiate normal and damaged cartilage. In the first step of determining the threshold value, Gaussian curve fitting was performed to fit the histogram of the CW. Normalized root mean square error (NRMSE) was used to determine the goodness of fit (normalizing against the variance of the target Gaussian curve). An average NRMSE value of 0.0187 ± 0.0189 with a maximum NRMSE value of 0.0553 was achieved during the fitting. Using the best fit Gaussian curve, threshold values for each dataset were determined. Using the threshold values derived using γ value of 95%, average percentage of damaged cartilage accurately identified (sensitivity) was $83.0 \pm 15.6\%$. On the other hand, the average percentage of normal cartilage accurately identified (specificity) was $97.9 \pm 0.9\%$. This shows that, using this technique, we were able to identify most of the damaged part of the cartilage while classifying only a very

small portion of the normal part of the cartilage as damaged (false positive) by using a γ value of 95%.

Fig. 2.5. Cumulative plot of normal part of the cartilage (in green) and damaged part of the cartilage (in red) derived from the 9 datasets.

The cartilage was plotted using surface mesh plot, as shown in Fig. 2.6. All the data were successfully visualized using the interface developed. In this study, each slice of the cartilage was divided into 20 regions and the average pixel value in each section was calculated and plotted. Thus the whole cartilage has $20 \times n$ sections, where n is the number of slices the cartilage covers. Results show that the damaged cartilage can be clearly visualized for all the dataset studied. The estimated volume per section for all the dataset is approximately 10 mm^3 . Although having more and smaller regions makes the cartilage more refine, the undesired effects caused by noise may become more prominent.

On the other hand, having fewer but larger regions may reduce the accuracy of the size of damage cartilage visualized. For a balance between noise removal and accuracy, the number of regions per slice could potentially be interactively varied by the radiologists during the examination, in our study, using 20 regions per slice worked well for the datasets analyzed.

Fig. 2.6. 3D plots of the cartilage using different values of γ . (a) $\gamma = 90\%$ - The 3D plot shows the lesion but together with relatively higher false positive. (b) $\gamma = 95\%$ - The 3D plot shows the lesion with negligible false positive result. (c) $\gamma = 99\%$ - The 3D plot shows only the most severe part of the lesion.

The threshold value was used to calibrate the color bar. Values below threshold are shown as blue and values above threshold span across all other colors. Fig. 2.6 shows the result of the 3D cartilage plot of one dataset using different threshold values to visualize the damaged part of the cartilage. The results show that only the more severe part of the lesion will be highlighted when a higher γ value (e.g. 99%) was used, whereas relatively more false positive were present when a smaller γ value (e.g. 90%) was used. However, the size of the false positive tends to very small and scattered which are less clinically significant.

Using the interface developed, it is possible to visualize the damaged portions of the cartilage in 3D and to interactively view the corresponding MR image to examine the location of the damage regions in relation to other anatomical structures (Fig. 2.7). It can be performed by clicking on the 3D plot where the clinician is interested to study. This function enables the radiologist to readily view the MR image to study the other details and verify the damage. This has been viewed by collaborating radiologists as a useful feature to have during the study of 3D models and MR images.

Fig. 2.7. A cartilage 3D plot together with two MR images showing damaged cartilage are shown, with a yellow cross indicating the region clicked on the 3D plot.

2.1.3. Discussions and Conclusion

This section (2.1) presents a methodology to quantitatively visualize focal cartilage damage of the femoral articular cartilage using clinically used MR PD weighted images. In this study, PD weighted images were used to study damaged part of the femoral articular cartilage. This type of images provides better tissue contrast with the surrounding tissues (e.g. meniscus and ligaments) and enables other surrounding tissues to be visualized, as compared to SPGR images. Although SPGR images have shown to be useful in developing automated segmentation techniques due to well defined cartilage boundary [66], radiologist generally prefers to use PD weighted images because of the need to study other structures together with the cartilage. The use of SPGR images has enabled the development of automated segmentation techniques for the cartilage [66] whereas the use of PD images in automated segmentation remains to be developed. Because our aim is to study whether it is feasible to visualize damaged cartilage in 3D using PD weighted images, the cartilage in all the datasets were manually segmented. PD images are less used in automatic image analysis because of its challenges in automatic segmentation. However, an indispensable advantage of using PD images directly in cartilage damage analysis is that this will aid clinicians to relate the locations of the cartilage damage directly to other periarticular and intraarticular structures using one set of images. This could potentially reduce the sequences required, hence, reducing scanning time. Furthermore, using PD weighted images directly may allow other relevant anatomical structures (e.g. bone marrow, meniscus and ligaments) to be extracted because they can be clearly seen in the images [137].

In our study, both the damaged and normal cartilage were segmented as a whole and used in the visualization process. This approach was chosen because it enables us to visualize the damaged cartilage based on the pixel intensities differences and we can estimate the volume of the damaged cartilage. The more commonly reported approach is to segment only the normal part of the cartilage and measure the remaining cartilage thickness [99, 128, 138-139]. In earlier reported studies, damaged part of the cartilage can be determined by visualizing regions of the cartilage where the thickness of the cartilage is abnormally thin, as compared to the rest of the cartilage. However, the volume of damaged cartilage cannot be determined using these techniques. A work by Losch et al. took a similar approach in which the damaged cartilage was estimated using SPGR images [138]. The main difficulty with using PD weighted image for cartilage damage analysis is that the contrast of the cartilage is not as well defined as that from fat suppressed SPGR images. Nonetheless, the cartilage could still be segmented through careful examination. An assumption made during the segmentation step is that the radiologist, based on his prior experience, could segment the cartilage as how he perceives the cartilage will be like when it is normal. Because our focus is on focal cartilage damage, instead of diffusive cartilage damage, the radiologist generally only needs to estimate an outline to include the damaged cartilage with a reference to the surrounding normal cartilage. This step is generally achievable by radiologists. Although the resultant area/volume of damage may be difficult to quantify reproducibly, it still provides a fair and usable estimate of the amount of cartilage damage/loss.

Our aim in this study is to investigate whether it is feasible to better visualize the damaged part of the cartilage using the differences in the pixel intensities between the normal part of the cartilage and the damaged part of the cartilage. The results from the study show that there is a significant difference in the pixel intensities between the normal part of the cartilage and the damaged part of the cartilage. Consequently, a Gaussian curve fitting technique was developed to determine the threshold value that can be used to differentiate the normal and damaged cartilage. The Gaussian curve fitting has shown to fit well to the main portion of the histogram which is mainly made up of the normal cartilage. By varying the percentile area γ under the Gaussian curve (e.g. from 90% - 99%), the presence of false positive can be controlled. Using a larger γ value (e.g. 99%), fewer number of pixels which belong to the normal part of the cartilage will be considered as damaged; and the number of damaged part of the cartilage that will be considered as normal will increase. Consequently, there is a tradeoff when determining this value. In our study, a γ value of 95% has shown to work well – achieving a sensitivity of 83.0% and specificity of 97.9%. It was observed that when a larger γ value is chosen (e.g. 99%), only the more severely damaged part of the cartilage was highlighted. Using a larger γ value also reduces the number of false positive but it will reduce the sensitivity. This γ value could potentially be interactively varied by the radiologists during the examination for different patients.

Our study demonstrates a way to detect cartilage lesion in PD weighted image. It complements to the technique by Lee et al. [125], which detects cartilage lesion based on cartilage thickness map from SPGR images. With the information of cartilage lesion

detection based on intensity of clinical PD image and the information based on cartilage thickness from SPGR image, the accuracy of detection may be increased.

Our approach allows the damaged portion of the cartilage to be visualized based on the pixel intensities differences. An important interactive feature developed for the interface is a 3D Trackback function which allows the clinician to study the 3D model of the cartilage and retrieve the corresponding MR image at the point of interest by clicking on the 3D cartilage model. Using the proposed visualization method, it is possible for clinicians to study anatomical details using MR images and reference the location of the cartilage with respect to anatomical landmarks of interest efficiently.

A technique to visualize cartilage damage is presented in this section. Using clinical PD weighted images and the visualization technique developed, it is possible to visualize the highlighted focal damaged cartilage in 3D. This could potentially provide a means for radiologist to better visualize the damaged cartilage and to report the damage quantitatively. In this study, because the focus is on visualizing damaged part of the cartilage and not differentiating patients who has normal cartilage and patient who has damaged cartilage, only MR datasets that have damaged cartilage (as diagnosed by radiologist) have been studied. Further, because the focus in this study is not on automating segmentation, manual segmentation was used to segment all the cartilage from the MR images. In the future, automated segmentation techniques can be developed to exploit the visualization technique presented in this section.

2.2. Characterizing Healthy and Damaged Cartilage in PD Images Using Histogram-based Parameters¹

Using PD images, clinicians are able to identify damaged part of the cartilage by spotting abnormal high signal presence and a change in cartilage morphology due to cartilage loss. This abnormally high signal intensity presence when the cartilage is damaged is generally due to the intrusion of fluid into the cartilage matrix. PD images are routinely used because it also gives excellent contrasts to other periarticular and intraarticular structures which are also important in the diagnosis of cartilage damage [130, 140].

In addition to the visualization framework developed in the previous section (2.1), this section (2.2) presents a study that performs statistical analysis of signal intensities of the cartilage using MR images and the percentage volume of damaged cartilage presence. The aim of the study is to investigate the characteristic of cartilage that has damage/lesions in a quantitative manner using histogram-based parameters. Because damaged cartilage tends to have abnormally high signal intensities than that of normal cartilage in PD images, we hypothesize that there is a relationship between the signal intensities of the cartilage and the size of the damaged cartilage presence, in terms of volume. In this study, we focus on the femoral articular cartilage.

¹ Part of this section has been published in: C.L. Poh, T.K. Chuah, and K. Sheah. Differentiating healthy cartilage and damaged cartilage using magnetic resonance images in a quantitative manner. 2010. DICTA, Sydney, NSW. Permission of reprint granted in the copyright agreement.

2.2.1. Methods

Pre-processing and Segmentation of Cartilage

Image preprocessing and segmentation were done as stated in section 2.1.1 except that 3 more MRI volumes with damaged cartilage and 5 more MRI volumes from healthy subjects were added to the analysis. All the additional dataset are of the same type: PD weighted images within the resolution and slice thickness mentioned. Three different masks/regions of interest are obtained in the segmentation step. The masks consist of (i) the whole cartilage which includes normal and damaged cartilage, (ii) only normal cartilage and (iii) only damaged cartilage. For cartilage with no damage, only one mask will be produced. These masks are then used to extract the signal intensities to be analyzed.

Analysis

The signal intensities from the normal and damaged cartilage were analyzed to derive statistical parameters (e.g. mean, median, standard deviation etc) using the segmented images. The total volume of the cartilage and the volume of damaged cartilage were calculated. The percentage of damaged cartilage is the ratio of the volume of damaged cartilage over the total volume of the cartilage (%CD). For normal cartilage without any damage, this percentage will be zero. Because signal from damaged cartilage tends to be abnormally high, there could be a difference in the mean and median of the signal intensity distribution which can be exploited. Hence, the difference between the mean

and median (mean-median) is calculated and used as the parameter for investigation. Linear regression is performed to determine the relationship between the parameters studied.

2.2.2. Results

A total of twelve MR data sets with different degrees of cartilage damage (D1–D12) and five data sets of normal cartilage (N1–N5) were used in this study. Table 2.1 shows the characteristics of the data used. Femoral articular cartilage was manually segmented using PD images into normal and damaged cartilage and the MR signal intensities were analyzed. The mean pixel intensities of healthy cartilage and mean pixel intensities of cartilage with damage are 85 ± 23.8 and 87 ± 29.3 , respectively. Average volume of the cartilage studied is $8574 \pm 3610 \text{ mm}^3$ and the average volume of damaged cartilage is $520 \pm 309 \text{ mm}^3$.

Table 2.1. Characteristics of the data used.

Sample	Standard Deviation of the total cartilage intensity	Mean-median	Total volume of cartilage (mm ³)	Volume of damaged cartilage(mm ³)	Percentage of damaged cartilage (%)
Cartilage with damage					
D1	33.1	4	17306	1327	8
D2	33.0	2	10057	582	6
D3	37.5	4	12274	806	7
D4	29.8	3	8361	612	7
D5	31.2	2	4517	302	7
D6	27.6	1	6152	383	6
D7	27.7	1	5941	350	6
D8	26.4	0	7372	368	5
D9	24.4	0	10957	575	5
D10	31.4	0	11122	503	5
D11	24.1	-1	7363	271	4
D12	25.0	-1	11026	170	2
Normal Cartilage					
N1	23.4	-1	10246		
N2	26.0	-1	4918		
N3	25.1	-1	1572	NIL	NIL
N4	23.4	-1	8870		
N5	21.2	-1	7705		

Fig. 2.8 shows the plot of mean-median versus the percentage of damaged cartilage presence. Results show that there is a positive linear relationship between the difference in mean and median of the cartilage signals (mean-median) and the percentage of damaged cartilage presence ($R^2 = 0.799$, $p < 0.01$). This suggests that there is a direct relationship between mean-median and the volume of damaged cartilage presence which in turn signifies the degree of damage.

Fig. 2.8. Relationship between the difference in Mean and Median and the percentage damaged cartilage presence.

For all the cartilages with damage, there is a strong positive correlation between %CD and mean-median. When the two variables (mean-median and %CD) were compared against one and another, it was observed that the sign (positive or negative) of the integer value of mean-median seems to be directly related to the value of %CD. When the %CD is greater than 5%, the mean-median has a positive value (mean-median > 0) whereas when the %CD is less than 5%, the mean-median has a negative value (mean-median < 0). In other words, the results show that when the cartilage has minor or no damage, the sign of the difference in mean and median tends to be negative whereas when the cartilage has greater degree of damage, the sign tends to be positive.

2.2.3. Discussions and Conclusion

This section presents a study that investigates the relationship between the signal intensities of the cartilage and the cartilage volume. The data were analyzed after cartilage was manually segmented using PD images. PD images were used in the study because it is one of the most commonly used sequences in knee joint diagnosis and it provides excellent contrasts among the different anatomical structures which needs to be studied as well.

Using the segmented images, the signal intensities of the cartilage and percentage volume of damaged cartilage were analyzed. Preliminary results suggest that there is a positive linear correlation between the mean-median of the cartilage intensities and the percentage of damaged cartilage. In other words, when the size of damaged cartilage increases the difference between the mean and median will also increase.

By comparing the mean-median and percentage of damaged cartilage, the results seem to imply that there exist a threshold value that can be used to differentiate normal cartilage and cartilage with damage. However, this value may not be sensitive enough to detect cartilage with very minor damage but might be more suitable to detect cartilage with more severe damage. Validation has to be done before any threshold value is used in clinical practice.

The difference between mean and median of the signal intensity distribution is relatively small (between -1 and 4) which might be sensitive to fluctuations in the signal intensities.

The range is small mainly because the range of grayscale is 8 bits (maximum value of 255). Increasing this range might increase the value of this difference which in turn may reduce its sensitivity. Future work will involve increasing the sample sizes and incorporating other quantitative MRI parameters [27].

In summary, the results of this study show that there is a positive linear relationship between the difference in mean and median of the cartilage signals and the percentage of damaged cartilage presence. This linear relationship held for all samples with damaged cartilage in our study. The results also show that when the cartilage has minor or no damage, the sign of the difference in mean and median tends to be negative whereas when the cartilage has greater degree of damage, the sign tends to be positive. This result suggests that there could be significant relationship between the characteristics of the signal intensities and degree of damage which can be exploited to differentiate cartilage that is normal and cartilage that has damage quantitatively after segmentation has been performed.

Chapter Summary

This chapter has presented a cartilage visualization framework that provides highlighting of locations that potentially have cartilage lesion. The visualization framework displays the cartilage as a 3D surface and enables radiologist to quickly switch from the 3D view to the conventional sagittal 2D slice. It could improve the efficiency of radiologists in the detection of cartilage lesion. The chapter has also studied histogram quantity with respect

to the percentage of cartilage damage in the whole cartilage and found that there is positive linear relationship between the percentage of cartilage damage and the mean-median metric.

Chapter 3

Texture Analysis and Classification of Bone Marrow Lesion in Knee MRI²

Introduction

Other than cartilage lesion, indications in bone marrow can also be useful for diagnosis of musculoskeletal disorders. Bone marrow lesion (BML) is an indication appearing in many kinds of musculoskeletal diseases including OA. Developing an intelligent CAD system capable of detecting and locating bone marrow affected by BML could improve the efficiency of diagnosis, such as by providing a means to screen normal subjects. In this chapter we develop the classification module of the CAD system by studying the use of texture analysis in the detection of BML. This chapter consists of two major sections: section 3.1 presents a study of texture parameters (individually) potentially useful in distinguishing between bone marrow with and without BML; in section 3.2, after establishing that there are individual parameters that display separability between bone

² The work in this chapter is derived from: (i) T.K. Chuah, C.L. Poh, and K. Sheah. Quantitative texture analysis of MRI images for detection of cartilage-related bone marrow edema. 2011. IEEE EMBC, Boston, MA. Permission of reprint granted in the copyright agreement. (ii) T.K. Chuah, C.L. Poh, E. Van Reeth and K. Sheah, Texture analysis and classification of bone marrow lesion in knee MRI - data from the Osteoarthritis Initiative. Magnetic Resonance Imaging. In press. Permission of reprint granted by Elsevier.

marrow with and without BML, we study and demonstrate the classification of bone marrow with and without BML using a subset of the parameters, with the aim to classify subjects who are normal and subjects who have BML.

3.1. Texture Analysis of Bone Marrow Lesion in Knee MRI

MRI is an imaging modality that can be used to visualize BML without the use of contrast agent, which is currently not possible by other modalities (including X-ray, computed tomography and ultrasound). Using fluid sensitive MRI sequences, BML appears hyperintense while surrounding unaffected marrow appears hypointense [141]. This appearance of BML in MRI is due to bone marrow fat being replaced by material containing H^+ ions [47]. Hence, BML is presented as free water signal [142]. Furthermore, MRI produces volumetric dataset which allows multi-planar analysis of the BML.

Using MRI, BML can be assessed either quantitatively [48], or semi-quantitatively using scoring system [49]. In [48], Mayerhoefer et al. demonstrated that the volume of the lesion can be quantified in a reproducible manner using computer-aided analysis. However, to carry out the quantitative assessment, segmentation of the bone marrow from image is necessary in which it was performed semi-automatically in [48] by applying threshold. Recently, Dijkstra et al. [92] reported a semi-automated segmentation method to address the problem of segmenting BML. All these assessments provide the severity of the BML, but the presence and the location of BML were not being detected automatically. Radiologists examine each and every slice to look for abnormality and/or

BML. The process is manual and subjective, hence carries the risk of inter-observer inconsistency. To improve the efficiency of analyzing the MR images and enable more consistent diagnosis, a computer aided diagnosis system capable of differentiating bone marrow with and without BML in MR images in a quantitatively manner is desired. To attain such a system, there is a need to first identify features that could be used to differentiate and characterize bone marrow with and without BML in knee MR images.

One approach to identify these features is through texture analysis (TA) of MR images. TA is a quantitative, computer-aided method that enables one to calculate image based mathematical patterns (of texture features) that can be used to differentiate and characterize different tissues [111]. TA has shown to be a promising technique in classifying pathological tissues from normal tissue [143-146]. Characterization of bones using MRI by TA (mainly using histomorphometric texture measures) has been studied fairly extensively, especially in the characterization of osteoporosis, as discussed in paragraph 3 of section 1.4.2. In the study of bone marrow disorder, which is one focus of our study, there is a need for texture parameters that are able to provide useful information about the presence of BML. To identify such texture parameters, we investigated the usefulness of textures in MRI data that indicate the presence of bone marrow lesion. These parameters could also complement the knowledge about pattern of the trabecular bones.

In this study (section 3.1), we investigated the possibility and ability of two-dimensional MRI based TA to characterize bone marrow that has and does not have BML, which

could then be used to detect the presence of BML in a quantitative manner. TA was performed on knee MR images and texture features obtained from a BML affected group and a normal control group were statistically compared. Ten most discriminative texture features were identified using feature selection methods. We focused on the bone marrow at the weight-bearing region of the distal femur since it is commonly affected by BML and of clinical importance. To the best of our knowledge, application of MRI based TA to characterize bone marrow with and without BML has not been reported.

3.1.1. Methods

MRI Data

Data analyzed in this study were obtained from the Osteoarthritis Initiative (OAI) database, which is available for public access at <http://www.oai.ucsf.edu/>. The specific datasets used are public-use data sets 0.2.2. OAI is a project collecting clinical, biological and imaging data of osteoarthritis patients to study the progression and disease pattern of osteoarthritis. One of the objectives in OAI is to identify imaging biomarkers (distinguishing texture features), which is addressed by our study. At the outset of our study, 29 knee MRI 3D images of subjects with osteoarthritis-related BML (identified by OAI, radiologically) were selected. We selected only subjects without signs of previous knee surgery such as anterior cruciate ligament replacement. The sizes of the lesions range from 240 mm³ to 17724 mm³ (1606 ± 3204 mm³), as reported by OAI. Same

number of control subjects who are not exposed to OA were selected matching the gender, age and race of the affected subjects. We excluded subjects that have non-OA related BML at weight-bearing region. As a result, a total of 58 volumetric knee images were analyzed. The purpose of matching is to have the same gender, age and race composition between the control and cases. Side of knee (left/right) was matched exactly. The type of magnetic resonance (MR) sequence studied is a sagittal intermediate-weighted turbo spin echo fat-suppressed (SAG IW TSE FS, TR = 3200 msec, TE = 30 msec) sequence, which is deemed as a suitable sequence to visualize BML. Using this MR sequence, BML appears to be bright patch with ill-defined boundaries, surrounded by the healthy bone marrow tissues that appear relatively dark. All data in this study are from OAI baseline visit. All images have the same spatial resolution: in-plane pixel size = 0.357×0.357 mm; in-plane matrix size = 444×448 ; slice thickness = 3 mm; and slice spacing = 3 mm.

Preprocessing of Images

The original gray-level images were first processed with brightness and contrast adjustment using ImageJ [134]. The brightness and contrast were adjusted automatically based on the middle slice (between lateral and medial) so that all gray tones of the image histogram of the whole knee span through all available gray levels. Each volumetric image was then processed using N4ITK bias correction algorithm to remove MRI bias field and low frequency variation in intensity [147], which could affect TA. The resulting bias corrected images were then used for TA.

Segmentation of Weight-bearing Region of Distal Femur – The Region of Interest

Knee MR images are viewed in the orientation of right knee, all left knees were mirrored to match the orientation of right knee for comparability. The bone marrow of all femur bones in the images were segmented by adopting a program by [98] which allows user to interactively define a set of points to outline the desired region, and shown to work well for segmentation using MR images. Cubic spline smoothing was then performed to produce a smooth contour from the set of user-specified points. We use this segmentation method as an improved equivalent of manual segmentation. The interactive segmentation produced a segmented mask of bone marrow for each femur.

The segmented mask obtained then consists of the shaft of the femur. However, because the shaft of the femur is not the weight-bearing area of the knee joint, any increased intensity in the marrow at that region is not considered as abnormal signal alteration. We only consider that an increased intensity in the marrow region near the femoral condyles indicates abnormality or presence of BML. Therefore, the shaft was removed from analysis. The shaft was systematically removed by first defining a cutting plane using 3 landmarks: the most proximal tip of cartilage at the lateral peak at the anterior portion of femoral condyle; and the most proximal tips of cartilage at the posterior portion of femoral condyles, one at each condyle. The mask of femur more proximal to this plane was then removed to form the mask of bone marrow at weight-bearing region. This resultant mask was used to define the region of interest (ROI) in the bias corrected image during TA. All segmented masks were verified by an experienced radiologist before

performing TA. A total of 1503 images were segmented. Representative examples of segmentations are shown in Fig. 3.1.

Fig. 3.1. Representative examples of MR knee images and segmentations. 1st row: a slice from a normal subject; 2nd row: an affected slice at the lateral femoral condyle; 3rd row: an affected slice at the medial femoral condyle. The regions of BML are indicated by solid arrows. 1st column: slice without segmentation; 2nd column: region of bone marrow delineated after initial segmentation (white outline); 3rd column: weight-bearing portion of the marrow delineated after the complete segmentation process (white outline).

Grouping and Characterization of Images

Slices that contain the femur bone marrow from all subjects were analyzed. They were separated into 2 groups: **Normal** and **Affected**. The Normal group consists of all the slices from control subjects and slices from *subjects* with BML but with no indication of BML at the particular slices; whereas the Affected group consists of all slices from subjects with BML and with the presence of BML. The presences of BML in these slices were verified by a radiologist at the outset of the study. There was a total of 1218 slices in the Normal group; and a total of 285 slices in the Affected group (regardless of the degree or area affected by BML). The difference in texture parameters at the weight-bearing area of the femur bone marrow between these 2 groups was studied.

Texture Analysis and Preliminary Feature Selection

TA was carried out using MaZda software version 4.6.2.0 [148], which provides the calculation of texture features. Analysis was done in 2D basis, slice by slice, due to the anisotropic nature of the data. Texture features were calculated for each ROI. These parameters include texture features based on gray-level histogram, co-occurrence matrix, run-length matrix, absolute gradient, autoregressive model and wavelet [111]. Texture parameters based on grey-level histogram characterize the appearance of a region of interest pixel-wise and does not take into account the arrangement of pixels, they are statistical summary of the grey level composition. Parameters based on co-occurrence matrix characterize a region based on the occurrence of pairs of gray values at a specified direction and distance. Texture parameters based on run-length matrix are summarizing quantities of the run lengths of different grey-levels in a region. Parameters based on

absolute gradient are mostly similar to those based on grey-level histogram but they are derived from the gradient image. Parameters from autoregressive model are derived under the assumption that there is local interaction between pixels and that a pixel is a weighted sum of its neighboring pixels, the weights and the noise component form the parameters. Texture parameters obtained from wavelet analysis are wavelet energies for specified regions after performing wavelet analysis, i.e. applying high pass or low pass filter and/or subsampling. The detailed definitions and formulas of texture features can be found in MaZda user's manual [149].

These 6 texture types were analyzed because they are deemed to be potentially capable of detecting the presence of BML: other than the texture parameters based on grey-level histogram, texture types listed here have been used in distinguishing bone marrow and fat [112]. Textures based on grey-level histogram were also included in this study because there is observed differences in grey levels between lesion and healthy marrow. For the co-occurrence matrix, the parameters were calculated in two distances ($d = 1$ and $d = 5$ pixels) because we have tested that for the type of images we are studying, any values of d in this range will merely give texture parameters that are highly correlated with each other, and also correlate with those of $d = 1$ and $d = 5$. Parameters of other texture types were all computed. As a result, the output of the TA is a vector of 147 texture parameters.

Individual Feature Selection

To test the significance of each texture feature in separating the Normal and Affected slices, statistical analyses were performed for the 147 texture parameters individually

(multidimensional analysis is covered in Section 3.2 instead). Differences in the texture feature between the 2 groups were analyzed using Mann-Whitney U -test (rank sum test). A p -value of less than 0.05 was considered statistically significant. This is performed using Matlab (The MathWorks Inc., Natick, MA, 2000). Non-parametric test was used because the data were ordinal in nature and do not have consistent distributional pattern.

To select top ten texture features that have the highest discriminative power for separation and classification, we used a number of commonly used feature selection methods. These methods include 1) Fisher coefficient; 2) Mahalanobis distance; 3) squared Euclidean distance; 4) 1-nearest neighbour leave-one-out classification performance (performance = 1 - error). These methods were performed using PRTools [150]. The features were then ranked according to the criteria of the methods, in decreasing significance.

Fisher coefficient which is defined as the ratio of between-class variance to within-class variance [151] which generally identifies highly discriminant features. Mahalanobis distance is a scale invariant multivariate effect size [152], which is a generalization of Euclidean distance that considers the distribution of data and the correlations of the dataset. It could be a better measure than Euclidean distance for data that are distributed non-spherically. The features were also selected based on squared Euclidean distance because it is the most intuitive and common way of measuring distances between classes. The performance of 1-nearest neighbor was also used as a measure to select features because it is the most basic machine learning algorithm.

3.1.2. Results

Significant Textures by Mann-Whitney U-test

We analyzed a total of 147 texture parameters to determine the number of parameters that differed statistically between affected slices (bone marrow with BML) and normal slices (bone marrow without BML). Table 3.1 summarizes the number of textural parameters which are significantly different between bone marrow with BML and bone marrow without BML ($p < 0.05$). There are 98 significant parameters from all 147 parameters analyzed, indicating that there exists prominent difference between the texture of affected marrow and normal marrow. This result suggests that the difference between the affected marrow and normal marrow can be represented quantitatively by image texture.

Table 3.1. Number of parameters from different textural types having statistically significant difference between affected image slice and normal image slice analyzed with Mann-Whitney U-test.

Texture Parameter Types	No. of Parameters Tested	No. of Parameters having Significant Difference
Histogram	9	7
Co-occurrence Matrix	88	56
Run-length Matrix	20	20
Absolute Gradient	5	3
Autoregressive Model	5	4
Wavelet	20	8
Total	147	98

Most Discriminating Textures

Table 3.2 lists the ten most discriminating texture parameters selected by different feature selection methods along with their p -values from corresponding Mann-Whitney U -test. (The detailed definitions of texture representations for the list can be found in MaZda user's manual [149]). To visualize the separability of the features, histograms of selected features were plotted.

Table 3.2. Ten most discriminating textures selected by different methods, and their corresponding p -values.

Ranking	Fisher		Mahalanobis		Euclidean		Nearest Neighbour	
	Parameter	p -value	Parameter	p -value	Parameter	p -value	Parameter	p -value
1	S(0,5)Correlat	<0.001	S(0,5)Correlat	<0.001	135dr_RLNonUni	<0.001	Perc.99%	<0.001
2	S(5,-5)Correlat	<0.001	S(5,-5)Correlat	<0.001	45dgr_RLNonUni	<0.001	Perc.90%	0.237
3	S(5,5)Correlat	<0.001	S(5,5)Correlat	<0.001	Horzl_RLNonUni	<0.001	Perc.01%	<0.001
4	S(5,0)Correlat	<0.001	S(5,0)Correlat	<0.001	Vertl_RLNonUni	<0.001	Perc.10%	<0.001
5	S(1,1)Correlat	<0.001	S(1,1)Correlat	<0.001	135dr_GLevNonU	<0.001	Perc.50%	<0.001
6	S(1,-1)Correlat	<0.001	S(1,-1)Correlat	<0.001	45dgr_GLevNonU	<0.001	S(5,-5)Correlat	<0.001
7	Sigma	<0.001	Sigma	<0.001	Horzl_GLevNonU	<0.001	S(5,0)Correlat	<0.001
8	S(0,1)Correlat	<0.001	S(0,1)Correlat	<0.001	WavEnLL_s-5	0.002	S(0,5)Correlat	<0.001
9	S(1,0)Correlat	<0.001	S(1,0)Correlat	<0.001	Vertl_GLevNonU	<0.001	S(0,1)Correlat	<0.001
10	Teta4	<0.001	Teta4	<0.001	Variance	<0.001	S(1,0)Correlat	<0.001

Correlat = correlation; Sigma = noise component in the autoregressive model; Teta = parameter in the autoregressive model; 135dr_RLNonUni = 135° run-length nonuniformity; 45dgr_RLNonUni = 45° run-length nonuniformity; Horzl = horizontal; Vertl = vertical; GLevNonU = grey-level nonuniformity; WavEnLL_s-5 = energy of wavelet coefficients in subband LL and scale 5; Perc. 99% = 99th Percentile.

Fig. 3.2 shows the histograms of the four most discriminative texture parameters selected by Fisher coefficient and Mahalanobis distance (both methods selected the same features). Each histogram is plotted with the number of bins equal to the nearest round number of $\sqrt{N}$, where N is the sample size for the particular group. It can be seen that for these four parameters, the histograms of Normal and Affected groups do not overlap much. This indicates that the rankings of these parameters are accurate, and suggests that these

parameters can be effectively used as supplemental information for making diagnostic decision or used for developing a decision boundary for multivariate classification system.

Fig. 3.2. Histograms of four most discriminative texture parameters selected using Fisher coefficient and Mahalanobis distance (the same 4 parameters for both methods).

Fig. 3.3 shows the histograms of four most discriminative texture parameters selected by Euclidean distance. For these parameters, the histograms of Normal and Affected groups have more overlapping. These parameters received higher ranks possibly because Euclidean distance is scale dependent. Note that the magnitudes of these parameters are much larger than other parameters. Fig. 3.4 shows the histograms of four most

discriminative texture parameters selected by the performance of 1-nearest neighbor classification. For these parameters, the histograms of Normal and Affected groups do not seem to separate at all, indeed, one of the parameter plotted here (Perc. 90%) failed to reject the null hypothesis in Mann-Whitney U -test ($p = 0.237$). These parameters turned out to be good in classification most likely because the instance at an adjacent slice (which is a dependent sample) in the same volume is determined as the nearest neighbor.

Fig. 3.3. Histograms of four most discriminative texture parameters selected using Euclidean distance.

Fig. 3.4. Histograms of four most discriminative texture parameters selected using the performance of 1-nearest neighbor classification.

3.1.3. Discussions and Conclusion

This section (3.1) has characterized bone marrow with BML in knee MRI using TA. Results show that a large number of feature parameters are potentially useful in discriminating affected slices from normal slices. There are parameters in all six types of textures having statistical significant difference between affected slices and normal slices. Seven out of the nine histogram-based features were statistically significantly different ($p < 0.001$). These indicate that the pixel intensity distribution is statistically significantly different between normal and affected subjects. This difference can be due to change in

the proportion of anatomical structure. The difference is translated into the difference in histogram-based parameters. Many of the co-occurrence matrix based feature were statistically significantly different ($p \leq 0.044$) between the two groups. The most important parameters are those of co-occurrence matrix based *correlation*. The *correlation* parameters were selected by three of the feature selection methods, confirming that they have good discriminative power. Note that the *correlation* parameters are important for all directions in the distance $d = 5$ (Table 3.2, Fisher, parameters of ranking 1 to 4), this agrees with the visual appearance that bone marrow does not have any directional orientation on the MR images. Moreover, run-length nonuniformity parameters in all directions were selected using Euclidean distance. Although this feature selection method is scale dependent, the selected parameters could still be effective classifiers in a multivariate classification system, due to their distance. Run-length nonuniformity is different in affected slices from that of normal slices possibly because the normal slices have more of uniform run distribution across intensities but the affected slices have more uneven distribution of runs across lower intensities and higher intensities. This is supported by the histograms in Fig. 3.3 because when runs are equally distributed across different grey levels, the parameter (run-length nonuniformity) takes on its lowest value [153].

Using the scale invariant feature selection methods, the most prominent textural difference between affected and normal slices is found in co-occurrence matrix based parameters, followed by parameters from autoregressive model. The results obtained in this study demonstrated the possibility of quantitative analysis to probe the existence of

BML. They suggest that there exist texture parameters that could aid in the differentiating bone marrow with BML and bone marrow without BML, as appeared in MRI. These features could in the future be incorporated into computer aided diagnosis system as a classification model to improve the diagnosis process: producing more consistent report based on quantitative image appearance instead of only visual perception, reducing risks and time associated with manual inspection thus improving efficiency. Having such a system will enable the presence and location of the BML to be determined in a more automated manner.

In this study, texture parameters that can potentially separate slices with and without BML were found, nonetheless, overlaps of textural histograms between Normal and Affected groups are still present in any parameters. This could be due to a number of reasons: (i) there is a spread of severity in the Affected group; (ii) the areas of healthy tissues in some affected slices dominate the textural values, which could potentially be overcome by 3D analysis; (iii) the texture value of some or part of the BML is actually similar to that of healthy tissues.

One limitation in this study is that we included only OA-related BML. The parameters found are believed to be useful for detecting BML regardless of its type. However, it does not provide any differential diagnosis to distinguish BML of different causes or diseases. The expertise of physicians, clinical data, patients' medical history and even laboratory findings are still needed to determine the exact type of BML one is having. It is also important to note that although the indications of BML across different slices are of

differing degree, the volumetric sizes of BML we have in our affected subjects have minimum size of 240 mm^3 , which is fairly large. It is unclear whether there will still be as many discriminative parameters if the size of edema is decreased. This leads to an intuitive future work, which is to perform TA with a larger size variation of BML. For this study, 2D image textures of BML were analyzed because most clinically available images are non-isotropic. Image textures in 3D, which is best analyzed with isotropic images, remain to be studied.

In conclusion, we have discovered the feasibility of TA in quantifying the difference between bone marrow with and without BML and found that the texture measures are potentially capable of differentiating between normal and affected slices. A number of texture parameters are significantly different between the two groups and this difference can be used in detection of BML, either by aiding radiologists directly, or by incorporation into a computer aided diagnosis system. The selected texture parameters also have potential to be used for localization and segmentation of BML.

3.2. Classification of Bone Marrow Lesion in Knee MRI by Texture Analysis

Section 3.1 discovered potentially useful parameters in separating slices with and without BML. This section (3.2) continues the investigation into selecting the feature set and testing the classification of slices and subjects to demonstrate the usefulness of texture-based analysis as part of the CAD system.

It has been shown that texture parameters can be used for to distinguish bone marrow and fat [112]. Therefore in this section, we hypothesize that the differences in the textural appearance of bone marrow with and without BML can effectively be used to distinguish and classify them. The objective is to demonstrate the feasibility of using texture parameters in the detection of BML, by studying the performance of texture-based classification systems that attempt to classify slices for BML and subsequently provide computer decisions of whether a subject has BML. In this study, textural parameters of MRI slices were computed and selected using one of the feature selection methods. Instead of individual feature selection as studied in section 3.1, the feature selection here (section 3.2) will be performed based on the multidimensional separation of the feature set because the focus of this section is not discovery of individual parameters but to demonstrate the classification using multiple parameters. (Note that in this section (3.1), the feature selection is individual selection so it does not take into account pair-wise or multivariate interaction between variables. The subset selection that considers multivariate interaction was done for classification in this section 3.2.) The slices were classified based on the selected textural features at the weight bearing region of the femur using different classifiers. In this thesis, the outcomes of eight different classifiers which train and test in reasonable time frame (< 5 minutes) were studied. Subsequently, the subjects were classified as normal or affected based on the slices classified by the best performing classifier.

3.2.1. Methods

Feature Selection for Classification

From section 3.1.1 (subheading Data Analysis), after performing the Mann Whitney U -test on the 147 textural parameters, 98 parameters were found to be significantly different ($p < 0.05$) between Normal and Affected slices. In this section (3.2), instead of selecting individual textural parameters independently from these 98 parameters, the textural parameters were selected as a subset, through forward feature selection (FFS) [154] based on inter-intra class distance (Fisher coefficient) [151], to avoid repeatedly selecting correlated parameters. This is important for classification task but not the concern when only attempting to discover individual useful parameters. The initial number of desired feature was arbitrarily set to 10.

To select the set of 10 optimum features exhaustively from the 98 significant features would involve prohibitively long time (there are 1.4×10^{13} combinations). FFS is a suboptimal feature selection strategy that provides a balance between performance and system complexity. It effectively reduces the computational time of exhaustive feature selection without compromising much of the accuracy of classification [154]. FFS first selects the best single feature (largest inter-intra distance), and then the best pair of features (based on multivariate inter-intra distance) is selected where the pair includes the best single feature; and the process continues by selecting one single feature at a time, that best paired with previously selected feature set, until the specified number of features is reached. Although the FFS is unable to drop a feature once it is selected, the advantage

of this characteristic is that it provides comparability for studying the optimum number of features (but not the optimum set of features, which is computationally prohibitive).

Classification of Slices and Subjects

Classification of slices was tested with different types of classifiers, and subsequently class labels for slices generated by the best performing classifier in the classification of slices will be used for classifying subjects. The 10 features selected using FFS were used for a binary classification of 1503 slices. To generate the class labels of each slice, all slices from each subject (a set of dependent samples) were isolated one at a time as testing set and the remaining data from 57 other subjects were used as training set to train a classifier. There were 58 cycles of training/testing until all slices were labeled. This process is similar to commonly used cross-validation procedure but it isolates a set of dependent slices for testing one at a time instead of isolating a set of randomly selected slices. The purpose of the process is applied to maximize the number of training sample and to simulate real world situation in which training samples are completely independent from the testing set. The classifications were performed using PRTools [150] and results of classification by eight commonly used classifiers were examined:

1. Normal densities based linear classifier (LDC)
2. Normal densities based quadratic classifier (QDC)
3. Mixture of Gaussians classifier, 2-mixtures (MOGC)
4. Minimum least square linear classifier (FISHERC)

5. Nearest mean scaled classifier (NMSC) assuming normal distributions with zero covariances and equal class variances.
6. Bayesian classifier (BAYESC), in which each feature axis is divided into N bins, N is optimized over the data range; number of training examples for each of the classes in each of the bins was counted, an object was classified to the class that gives maximum posterior probability.
7. K -nearest neighbor classifier (KNNC), K is optimized (hence a variable value for each training) by the leave-one-out error of the training data.
8. Levenberg-Marquardt trained feed-forward neural net classifier (LMNC), 6 units in one hidden layer (6 is the optimized number between 1 to 10 units with one hidden layer).

A good choice of classifier usually depends on the knowledge of the distribution and the pattern of data. However, the distributions of 10-dimensional data are difficult to visualize. Since the feature score distributions were unknown, different types of classifiers were examined for comparison. In this study, we included parametric classifiers (LDC, QDC), statistics based classifiers (FISHERC, NMSC, BAYESC and KNNC), classifier based on unsupervised learning MOGC, and non-linear classifier (LMNC).

The performances of the classifiers were evaluated with respect to their ability to classify image slices (affected by BML or normal). The best performing classifier (amongst the 8) according to area under the ROC (AUC) was then analyzed. AUC was calculated using

the true labels and the labels generated by each classifier (binary), with the positive instances being the instances with BML. To evaluate the reliability of AUC of slice classification, its p -value³ was calculated under the null hypothesis that the classification has no skill in distinguishing affected slices from normal slices. The equations for calculation are:

$$SE(Area) = \sqrt{\frac{A(1-A) + (n_A - 1)(Q_1 - A^2) + (n_N - 1)(Q_2 - A^2)}{n_A n_N}}$$

$$Q_1 = \frac{A}{2 - A}$$

$$Q_2 = \frac{2A^2}{1 + A}$$

Where A is area under ROC curve, SE is standard error, n_A is the number of affected slices and n_N is the number of normal slices.

The two-tailed p -value is then calculated based on the z ratio using normal distribution:

$$z = \frac{A - 0.5}{SE_{A=0.5}}.$$

For subject classification, the class labels generated for slices by the best performing classifier were used to determine whether a subject is affected by BML. To decide whether a subject is affected by BML, a threshold is set such that if the number of slices classified as affected reaches the threshold, the subject is classified to be affected. We studied the positive predictive value (PPV), negative predictive value (NPV) and AUC of classification of subjects with respect to different thresholds.

³ Concept and equations from (i) J.A. Hanley and B.J. McNeil, The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, **1982**. **143**(1): p. 29-36. (ii) GraphPad. Details of calculations for ROC curves. <http://graphpad.com/support/faq/details-of-calculations-for-roc-curves/> 16 Nov 2012.

3.2.2. Results

Feature Selection

Table 3.3 shows the ten textural parameters selected by FFS listed in decreasing importance. The most distinguishing feature is $S(0,5)Correlat$, which is the correlation in the vertical direction at a distance of 5 pixels. The parameter is based on co-occurrence matrix. (The detailed definitions of texture representations for the list can be found in MaZda user's manual [149].) As expected from FFS which maximizes its metric as it includes more features, the first 3 selected features are uncorrelated with each other: the 1st is based on co-occurrence matrix, the 2nd is based on run-length matrix and the 3rd is based on histogram. The top 3 features are less likely to be correlated with each other because they are of different textural types. It is confirmed here that including grey level histogram-based parameters is a reasonable choice because parameters such as skewness and kurtosis were selected as the top 10 parameters by FFS.

Table 3.3. Ten features selected by forward feature selection based on inter-intra class distance.

Importance	Feature
1st	S(0,5)Correlat
2nd	135dr_GLvNonU
3rd	Skewness
4th	Teta3
5th	Kurtosis
6th	Perc.01%
7th	GrVariance
8th	GrMean
9th	S(1,-1)SumEntrp
10th	S(5,5)Correlat

Correlat = correlation; 135dr_RLNonUni = 135° run-length nonuniformity; Teta = parameter in the autoregressive model; Perc. 01% = 1st Percentile; GrVariance = variance of absolute gradient; GrMean = mean absolute gradient; SumEntrp = sum entropy.

Classification of Slices

Table 3.4 shows the results of classifying slices based on the ten selected features and using eight different classifiers. The AUC were shown with its *p*-value under the null hypothesis that the classification has no skill in distinguishing affected slices from normal slices. The performances of the classifiers were compared using AUC because the data were unbalanced with respect to the instances for each class. Nearest mean scaled classifier (NMSC) and Quadratic classifier (QDC) gave the best classification result which the AUC is 0.780. The ROC curve of NMSC is plotted (Fig. 3.5) in comparison with ROC curves of other classifiers, and because it has the highest AUC and it is faster than QDC, this best classifier was used for studying the classification of subjects. Note that the ROC curve of each classifier is only defined by one point (which represents the lowest false positive rate possible while maximizing true positive rate) and two straight lines because each classifier used is a discrete classifier.

Table 3.4. AUC and p -values of the AUC using different classifiers.

Classifier	AUC	p -value
LDC	0.772	<0.001
QDC	0.780	<0.001
MOGC	0.758	<0.001
FISHERC	0.756	<0.001
NMSC	0.780	<0.001
BAYESC	0.762	<0.001
KNNC	0.502	0.906
LMNC	0.763	<0.001

Fig. 3.5. ROC curves of classifications of Normal and Affected slices using various classifiers.

Classification of Subjects

The subjects were classified as either affected or normal using the class labels of slices that were assigned by NMSC. The classification results are shown in Table 3.5. The table lists PPV, NPV and AUC with respect to different number of slices used as threshold to determine the disease status of a subject. It can be seen that using 5 slices as the threshold (need not be consecutive slices) gives the highest AUC, 0.914. The ROC curve of this case is shown in Fig. 3.6 (black) in comparison with the curves using different thresholds. Using a threshold of 5, the classification of subjects resulted in only one healthy subject (out of 29) and three affected subjects (out of 29) being misclassified.

Table 3.5. Classification of subjects using different number of slices as threshold.

Accuracy measures	Threshold (number of slices)						
	1	2	3	4	5	6	7
PPV	0.630	0.771	0.844	0.867	<i>0.962</i>	1.000	1.000
NPV	1.000	0.913	0.923	0.893	<i>0.875</i>	0.784	0.725
AUC	0.707	0.828	0.879	0.879	<i>0.914</i>	0.862	0.810

Fig. 3.6. ROC curves for classification of subjects using 1, 3, 5 and 7 slices as threshold.

3.2.3. Discussions and Conclusion

This study (section 3.2) demonstrated the feasibility of classifying MR image slices and subjects with and without BML, based on textural information. AUC of 0.780 obtained using NMSC implied that the classifier can classify a positive slice over a negative slice correctly nearly 80% of the time. Subsequent classification of subjects using 5 slices as threshold can classify an affected subject from normal subject correctly over 90% of the time. Comparing the performances of classifiers in classifying slices (Table 3.4 and Fig. 3.5), most of the turning points of ROC curves are concentrated at the left-hand side of

the plot. This suggests that texture based classifier classifies a slice as positive only with strong evidence and tends to have low false positive rates [155]. Contrarily, it is important to be aware of the false negatives produced by this kind of classifier. All classifiers studied have similar accuracy except that K -nearest neighbor classifier has very low AUC, its performance is as good as a random guess. This could be due to the scales of different features: the less important features that have larger scales took undesired dominance over the more discriminating ones.

The overall performance of subject classification was better than that of individual slice classification. This suggests that although the AUC of slice classification is not very high (0.780), the use of threshold in classifying subject can consolidate and partially remove error made in slice classification which could be due to too small an indication of BML in some slices (causing false negative), or some random signal alteration that simply resembles BML (causing false positive). The classification of slices using discrete classifier provides a way to classify subjects by thresholding the number of slices affected. The choice of this threshold can then be used for different purposes (Table 3.5). One can choose to maximize PPV or NPV in an application or to choose to use 5 slices as threshold to optimize between them. A low threshold such as 1, will compensate the shortcoming of the slice classifier in making false negatives, but inevitably made a number of false positives (subjects). For practical interest, when we choose 1 slice as threshold, the classifier would be correct all the times when it identifies a negative case ($NPV = 1$) while more than 60% of the positive cases are correctly identified. This could potentially half the time a radiologist needs to screen out healthy-looking subjects, and

can pay more attention to those more likely to be pathological. A high threshold would not be of practical interest because it tends to make false negatives, which means it can miss a positive diagnosis.

To confirm that the number of features initially used is indeed a good choice, the feature selection and classification cycle were re-run using NMSC and the AUCs for each run were plotted against number of features in Fig. 3.7 (feature selection was re-run only when more features need to be selected). It is clear that the initially chosen number of features, 10, is the best number of features chosen by FFS for NMSC classifier which gives the highest AUC value in the plot, suggesting that this number is suitable. Furthermore, the value of AUC started to decrease when the number of features was larger than 10, likely because of the sparseness of data the dimensionality has created (curse of dimensionality).

Fig. 3.7. AUC using NMSC with different number of features in the classification of Normal and Affected slices.

The classification presented here nevertheless has some limitations. We have examined the false cases and found out some exemplar false cases to explain the reason of error. Fig. 3.8 shows examples of false positive and false negative, along with their outlines of segmentations. It can be seen that the false positive case has relatively homogeneous appearance but there is some signal alteration (mainly increased signal, possibly due to partial volume effect or healthy variation of signal) at the core of the marrow. This may be the cause that the slice is identified as affected. Other false positives also have similar characteristics and mainly occurred at the extreme medial or extreme lateral slices. False negative usually occurs when there is very subtle indication of signal alteration such as

the example shown, that occurs at the distal femur tip. The error occurred because the surrounding healthy tissues overpowered (in terms of size) the textural value originated from the BML (not the osteophyte at the anterior tip). Since the subject classification is based on the number of slices classified as affected (threshold = 5), affected subjects were misclassified when the BML covers only very few slices, regardless of whether the lesion appears large or small in sagittal plane. This problem could potentially be addressed by multiplanar analysis that includes textures of other planes.

This classification method can be applied in CAD system to improve the accuracy and consistency of diagnosis by providing the radiologists a second opinion. The results of subject classification presented here is highly accurate. This affirms that textural information is indeed able to separate subjects affected by BML from normal ones. In addition to textural information from the images, a computer aided diagnosis system could utilize more information such as body mass index, age, or even the count of white blood cells to further enhance the performance of classification. However, segmentation of the bone marrow from the MRI is necessary before this can be an automatic process. In this study, we have used an equivalent of manual segmentation to delineate the femur bone marrow and it provided good representation of the region of interest. Automatic segmentation methods could be developed in the future to delineate the bone marrow region to automate the derivation of the texture features.

Fig. 3.8. Examples of erroneously identified slices using the 10-feature NMSC algorithm. 1st row: original slice and the segmentation of a false positive. There is no focus of subchondral BML. 2nd row: original slice and the segmentation of a false negative slice, where there is positive subchondral BML (arrow).

To conclude, the study in this section has demonstrated the feasibility of classifying slices and subjects with and without bone marrow lesion, reporting a satisfactory accuracy for slices classification and high accuracy for subject classification. It is unclear exactly what size BML would become clinically significant, nevertheless the challenge in the future will be to identify smaller BML and to classify different pathologies: osteophyte, bone tumor, transient BML and others.

Chapter Summary

This chapter has studied textural parameters potentially useful in the detection of BML and discovered a number of useful parameters. The chapter has also further the investigation in selection of feature subset and classification to demonstrate the usefulness of the textural parameters in classifying slices and subjects with and without BML, achieving an AUC of 0.914 for subjects classification. This reported accuracy of slice classification is very encouraging and the accuracy of subject classification is very high. This suggests that the textural parameters and the classification method applied, being part of the CAD system, could aid in the diagnosis process of musculoskeletal disorder.

Chapter 4

Advancements for Active Contour-based Automated Bone Segmentation

Introduction

The earlier study presented in Chapter 3 has shown that it is feasible to use texture analysis to classify between bone with and without BML. However, the segmentation process was still carried out almost manually which is highly laborious and repetitive. Consequently, it is highly desirable that the task of segmenting the bone be automated. This chapter presents studies performed to advance active contour segmentation algorithms for bone in knee MRI. Section 4.1 presents an advancement which provides an efficient and reproducible way to automatically terminate contour, applicable for all kinds of active contours. Section 4.2 presents the development of an initialization strategy made for the use of parametric active contour. The initialization strategy provides the contour a targeted direction for movement towards intended object boundary in knee bone segmentation on MRI.

4.1. Developing Termination Criterion for Active Contour⁴

Active contour is a component algorithm capable to delineating a targeted anatomical structure based on region and/or edge information, and it can be applied to segment bone in conjunction with a model. For both parametric and geodesic active contours, terminating the contour evolution in a fast and accurate manner has been a challenge that requires further research attention. Although theoretically the formulation of active contour expects that there will be no movement of the contour when it reaches the desired location (such as reaching a force balance of internal and external force), the ideal force balance and zero contour movement is practically difficult to obtain because it may involve any or all of these: very small time step, long evolution time, and extensive tuning of force parameters. In view of these difficulties, several termination criteria were proposed [156-161].

For parametric active contour, existing methods are mostly based on pixel movements of the contour. They include sum of norms [156], sum of distance, maximum distance, minimum distance, median distance and average distance. Ray et al. has also proposed a definite termination criterion that is based on the direction of pixel movements [159]. The criterion terminates the evolution when the percentage of contour pixels moving in opposing directions in subsequent iterations is above certain predetermined threshold. Generally, termination criteria used in parametric active contour are not applicable to geodesic active contour because there is no pixel correspondence between the curves of

⁴ Part of this section has been published in: T.K. Chuah, J.H. Lim, and C.L. Poh. Group Average Difference: A Termination Criterion for Active Contour. *Journal of Digital Imaging*, **2012**: **25**(2) p. 279-93. (Reprinted with permission from Springer, license number: 3060701258206).

subsequent iterations in geodesic active contour. For geodesic active contour, termination can be based on the energy functional [158, 160], entropy and simulated annealing [157], and the change of length of the contour [161]. The termination criteria used for geodesic active contour, such as that used in [158, 160] has not been demonstrated in the use of parametric active contour. Hence their applicability is still unclear. Despite much effort in proposing new termination criteria, studies that assess termination criteria by systematically comparing them are sparse.

To assess how well a termination criterion performs for both types of active contours, three important factors should be considered. They include accuracy, segmentation speed and shape insensitivity of the criterion. Shape insensitivity measures the extent that shape change of the target shape can have (while holding the termination parameters constant) before the criterion fails. A criterion is judged to be a failure if it does not return an accurate result within reasonable time or a number of iterations. Although shape insensitivity is a qualitative measure, it is important in practice because shape of target structure can vary significantly across different samples (of the same anatomical location) and/or different slices (in the same volumetric dataset). Different slice locations may need to be segmented at different instances or for different patients. An active contour termination criterion is useful in practice if its performance (such as segmentation time and accuracy) is insensitive to this shape variation [162].

One more factor should be considered for the assessment of termination criteria for parametric active contour: insensitivity to contour resampling. Contour resampling is the

addition and removal of contour vertices in subsequent iterations as suggested in [163]. The contour resampling step is necessary for avoiding vertex clustering and vertex dispersion. A termination criterion insensitive to contour resampling is applicable for both types of parametric active contour: with and without resampling. Contour resampling results in the loss of pixel correspondence of contours in subsequent iterations. As a result, all pixel movement based termination criteria are inapplicable to contour with resampling. They are considered to be sensitive to contour resampling.

This section presents a novel termination criterion called group average difference (GAD) which is designed for parametric Gradient Vector Flow (GVF) active contour [82] with contour resampling and unit normal force selection. This termination criterion is based on image difference between subsequent iterations, as opposed to most of the existing criteria that use pixel movement [156, 159]. This section describes studies that evaluate the proposed termination criterion in terms of segmentation speed, accuracy, shape insensitivity and insensitivity to contour resampling by comparing to existing termination criteria. The merits and limitations of the proposed criterion are also discussed.

4.1.1. Methods

This section (section 4.1) presents the GAD termination criterion by first studying its trend line and comparing with that of existing termination criteria (*trend line study*). Secondly, the termination criterion is evaluated with synthetic images of different target shapes (*time-accuracy study*). To demonstrate that the proposed criterion is applicable to

real medical images, we applied the GAD for the segmentations of three different anatomical structures of varying shapes using medical images. Sample anonymized medical images were used with ethics approval. The segmentations in medical images include femur, tibia and humerus using MR images. The anatomical structures chosen have varying shapes across different slices within a volumetric dataset. Segmentation of these structures is important for visualization in diagnosis and surgical planning [164-166]. All medical images used underwent Otsu's thresholding [167] before the application of the active contour so that the active contour essentially evolved on an edge map of a binary image.

The purpose of the *trend line study* is to compare the trend of the proposed criterion with the trends of existing criteria. This comparison enables one to select or design a steadily decreasing metric for termination without having to change the energy functional of the active contour. The comparison of trend lines also allows visualization of the robustness or weakness of a termination metric.

The *time-accuracy study* was carried out for studying the termination time, accuracy and shape sensitivity of existing criteria and the proposed criterion. In this study, we compare the average pixel movement (a.k.a. average distance) with GAD across six different test shapes of varied complexity using synthetic images. Average pixel movement (APM) is chosen amongst other existing termination criteria because it is representative to most other pixel movement based criteria as shown from the results of the *trend line study*. The shapes tested include various curvatures, corners, inward and outward concavities.

Both studies in this section employ the GVF active contour with contour resampling (no contour resampling for tests of APM termination method) and unit normal force selection.

GVF Snake with Contour Resampling and Unit Normal Force Selection

The parametric GVF snake used in this section follows the force balance equation [82]

$$\mathbf{x}_t(s, t) = \alpha \mathbf{x}''(s, t) - \beta \mathbf{x}'''(s, t) + \mathbf{v} \quad (4-1)$$

where s is the parameter of the contour $\mathbf{x} = [x(s), y(s)]$; x, y are the horizontal and vertical spatial coordinates respectively; vector field $\mathbf{v} = [u(x, y), v(x, y)]$ is the GVF field; α and β are weights for first and second derivative terms of internal forces in the energy functional of traditional snake [75]. Solution of Eq. 4-1 provides the critical point (minimum in this case) of the energy functional. Left hand side of the equation is the time derivative of the contour (subscript here indicates derivative); right hand side of the equation is the balanced force. Both the balanced force and the time derivative of the contour equal zero at the critical point.

The balanced force comprises both internal force and external force (GVF field) $\mathbf{v}$. The internal force is made up of second and forth partial derivatives of the contour with respect to the parameter s . For all the images processed in this section, we use $\alpha = 0.85$

and $\beta = 0.01$. Snake evolution time step ζ is set to 0.075. These values (α , β and ζ) were chosen based on results from a separate optimization study of the active contour. Forward time and central space numerical schemes were applied for implementation.

To derive $\mathbf{v}$, the edge map $f(x, y)$ used is

$$f(x,y) = |\nabla [G_\sigma(x,y) * I(x,y)]|^2 \quad (4-2)$$

$I(x, y)$ is the thresholded binary image of interest; $G_\sigma(x, y)$ is a 5×5 two-dimensional Gaussian kernel with a standard deviation $\sigma = 2$; ∇ is the gradient operator and $*$ is the convolution operator. The edge map is employed in the numerical solution for the GVF components u and v as follows:

$$u_t(x, y, t) = \mu \nabla^2 u(x, y, t) - [u(x, y, t) - f_x(x, y)] \cdot [f_x(x, y)^2 + f_y(x, y)^2] \quad (4-3a)$$

$$v_t(x, y, t) = \mu \nabla^2 v(x, y, t) - [v(x, y, t) - f_y(x, y)] \cdot [f_x(x, y)^2 + f_y(x, y)^2] \quad (4-3b)$$

The parameter μ is the regularization term associated with smoothing relative to image gradient; u_t and v_t are time derivatives of the horizontal and vertical GVF components respectively; f_x and f_y are horizontal and vertical spatial derivative of the edge map respectively; ∇^2 is the Laplacian operator.

The regularization term μ is set to 0.2. While solving Eq. 4-3a and 4-3b, u and v were initialized using the directions (having magnitudes of unity) of the external force field of the traditional snake. Rounded value of $\sqrt{N}$ iterations were applied (N is maximum of the image size along either horizontal or vertical direction). The resulting GVF from this solution is a static force field which does not change in the process of snake evolution. The iterations use a time step (ζ_{GVF}) value of 0.82 for the image used in *trend line study* and medical images in *time-accuracy study*, and a value of 0.01 for the synthetic shapes in *time-accuracy study*. This time step value is chosen based on the closeness between the initial contour and the target boundary. Whenever termination criteria are compared in this section, these evolution parameters mentioned are set to be the same in the criteria. (Note: not to confuse ζ_{GVF} , the time step for obtaining GVF with ζ which is the time step for snake evolution)

During contour evolution, all the external forces encountered by the contour are resolved into the tangential and normal directions with respect to the contour. The tangential component of each force is discarded because it does not contribute to the shape of the contour, but causes problem of vertex dispersion or vertex clustering of the contour. If normal component is present on the contour, only the direction (magnitude of unity) is used to step the contour forward in time. The process of incorporating normal force onto the contour for evolution is called *normal force selection*. Taking unity magnitude for normal force makes the contour moves in a more consistent way which is dominated by the external forces relative to the shape of the contour. To further avoid vertex dispersion and clustering problem, contour resampling operation is carried out (but not for

evolutions terminated based on pixel movement) for each time step of the evolving contour as suggested by [163].

Pixel Movement Based Termination Criteria

This subsection defines existing [156] (Eq. 4-4) and potential (Eq. 4-5 to 4-9) pixel movement based termination metrics for parametric active contour. A contour on a digital 2D image is defined by a finite number of pixels that represents the vertices of the contour. Let this finite number be n ; the vector $\bar{x}$ is a column vector that lists the horizontal coordinates and the vector $\bar{y}$ is the column vector that lists the vertical coordinates of the n points. We denote $\bar{x}$ and $\bar{y}$ at i -th iteration as $\bar{x}_i$ and $\bar{y}_i$. For contour without resampling during evolution, n is a fixed number throughout evolutions. Let the distance travelled by each of n contour pixels at i -th time step be

$$d_i = \left[(\bar{x}_i - \bar{x}_{i-1})^2 + (\bar{y}_i - \bar{y}_{i-1})^2 \right]^{\frac{1}{2}}$$

The termination metrics at i -th time step are

$$\text{Sum of norm} = \|\bar{x}_i - \bar{x}_{i-1}\| + \|\bar{y}_i - \bar{y}_{i-1}\| \quad (4-4)$$

$$\text{Sum of distance} = \sum_1^n d_i \quad (4-5)$$

$$\text{Average distance (APM)} = \frac{\sum_{i=1}^n d_i}{n} \quad (4-6)$$

$$\text{Maximum distance} = \max(d_i) \quad (4-7)$$

$$\text{Minimum distance} = \min(d_i) \quad (4-8)$$

$$\text{Median distance} = \text{median}(d_i) . \quad (4-9)$$

The symbol $\|\dots\|$ denotes Euclidean norm. The sum of distance is obtained by first calculating the distances moved by each of n contour pixels, and then all the n distances are added together to form the metric.

Group Average Difference Termination Criterion

Group average difference is a termination criterion based on image difference between subsequent iterations calculated using XOR logical operation. To calculate it, the image difference is obtained by summing up the number of pixels that are different in subsequent iteration. Afterwards, the group average image difference is taken over a predetermined number of iterations, for example, 50 iterations; and group average difference is calculated for instance, between the 1st and 2nd group of the (non-overlapping) 50 iterations. Fig. 4.1 shows an illustration of exemplar calculation and subsequent paragraphs provide detailed explanations and generalized formulation for the calculation.

Fig. 4.1. Illustration of exemplar calculation of GAD.

Once the contour evolves through one time step, a list of coordinates that represents the vertices of the contour in double precision is obtained. The coordinates are rounded when they are drawn onto the pixelized image. This rounding causes the contour to be broken at some instances. Fortunately due to the resampling operation, the breakage of the contour is no larger than one pixel at every instance. This breakage is repaired on the binary image describing the contour with the background using the following method:

- 1) apply Gaussian lowpass filter (3×3 kernel size, $\sigma = 0.5$) after changing the binary image to double precision;
- 2) convert the double precision image back to binary image, with all non-zero pixels eventually carries the value of 1;
- 3) thin the contour to one pixel width using the morphological algorithm described in [168].

Once the contour has been repaired, the binary image is filled with a morphological hole-filling operation. This operation fills any background pixels that cannot be reached by filling the background from the edge of the image. The filling is based on 4-connectivity

neighborhood. A filled area is obtained for every iteration step. This filled area forms the segmented mask of target structure and is the basis for setting termination criterion.

Two images of filled areas in subsequent iterations are compared using XOR logical operator. The more the two images share the same filled area, the less logical ones will appear in the resulting image. Denoting the filled image of i -th iteration as I_i , for an $M \times N$ binary image (M is the size at the vertical direction and N is the size at the horizontal direction) we have

$$I_i \text{ XOR } I_{i-1} = I_d(i), \quad 2 \leq i < \infty \quad (4-10)$$

$$[D]_k = \left[\sum_{y=1}^M \sum_{x=1}^N I_d(x, y) \right]_i, \quad k = i-1 \quad (4-11)$$

By summing up the number of logical ones in I_d , we would expect a decreasing metric as the snake approaches equilibrium position. The quantity D is denoted as the image difference. The index k is the iteration index of D which is always one integer less than i .

However, due to the use of only normal force of unity magnitude on the contour for evolution, D will not decrease to zero. This is because when the contour is near the equilibrium position, the contour location will fluctuate within 3 pixels around the exact edge. This is a pseudo-equilibrium state. The fluctuation will only be of practical concern if the image is small (e.g. smaller than 100×100 pixels) and/or the target structure is thin (e.g. having thickness smaller than 10 pixels). Recognizing that when all the vertices of

the contour reach their pseudo-equilibrium position, the value D will fluctuate around a constant, we define the quantity called the *group average*, G .

$$[G]_j = \frac{\sum_{k=(1+g(j-1))}^{jg} D_k}{g}, \quad 1 \leq j < \infty, \quad g \in \mathbf{Z} \quad (4-12)$$

In Eq. (4-12), g is an integer representing the number of iterations to form a group. The group average has iteration index j . The calculation of group average is non-overlapping (i.e. the data used for computing the first group average is not used for computing the second group average etc). Since D fluctuates around a constant, we can expect G to be approaching a constant when g is reasonably large. The value of g is selected to be 200 in our application. For a quantity that approaches zero while the contour progresses and settles, we define the *group average difference* (E) as

$$[E]_p = [G_{j+1} - G_j], \quad p = j, \quad 1 \leq p < \infty, \quad (4-13)$$

The index p is the iteration index of the group average difference. The evolving contour is expected to terminate when p is increasing and E is approaching zero.

Analysis

For the *trend line study*, we study the trend line of each metric versus the number of iterations, up to 1200 iterations (equivalent to 1200 time steps). The decrease of trend

and steadiness of decrease are examined. A steadily decreasing metric is deemed to be a metric more suitable for practical applications. It is because a steadily decreasing metric provides the ease of determining a threshold for termination, i.e. a value near to zero; whereas it is difficult to determine a threshold for a non-decreasing metric.

For both the *trend line study* and the *time-accuracy study*, the maximum number of iterations is set at 10 000. If a termination metric threshold cannot be reached within this maximum number of iterations, the evolution is forced terminated.

In the *time-accuracy study*, the time to termination is used as an indicator of speed of segmentation. All contour evolutions were implemented using MATLAB (The MathWorks, Inc.) on IBM workstation with Intel Xeon CPU 5160 @ 3.00GHz. Number of iterations to termination is used as an indicator to shape sensitivity of a termination method. Failure to terminate within 10 000 iterations is deemed as the failure of a particular termination criterion to terminate within reasonable time. Shape sensitivity is measured by the frequency of failure of a termination metric to terminate within reasonable time. By having unified threshold values for APM and GAD, the shape sensitivity of the termination criteria can be studied.

For all the segmentations carried out using active contour, accuracy is assessed using dice similarity coefficient (DSC) [71], verified with the ground truth. The ground truths for medical images are obtained from manual segmentation carried out by expert. This coefficient is a measure of overlap between the ground truth and the active contour

segmentation. It takes into account both positive overlap and negative overlap because it decreases as area of false positive and/or false negative increase(s).

4.1.2. Results

Trend Line Study

In this study, trend lines of termination metrics of the active contour on sixteen images (six synthetic images and ten real medical images) were analyzed. The six synthetic images comprise binary images of different shapes, and the real medical images are MR images of femur, tibia and humerus. In general, the results show that it is difficult to find a good threshold for all shapes using pixel movement based metrics, whereas GAD shows steadily decreasing trend approaching zero for all shapes which enables threshold to be better determined. For simplicity, trend lines of one representative image are presented. The representative image is a knee magnetic resonance imaging (MRI) slice with a size of 171×171 pixels.

The initial contour for this study is shown in Fig. 4.2a. The final contour for pixel movement based evolution (without contour resampling each time step) is shown in Fig. 4.2b; and the final contour for evolution with contour resampling (that is used in obtaining GAD) is shown in Fig. 4.2c. All of the contours are shown overlap with the MRI slice used (histogram equalized for display). To show an example of how the pixel

movement based criteria vary over evolution time steps, the six termination metrics have been plotted in Fig. 4.3 using the representative knee MR image.

Fig. 4.2. Comparison of segmentation with pixel movement based evolution with evolution with contour resampling. (a) Initial contour on the MRI slice used. (b) Final contour from pixel movement based evolution. (c) Final contour for evolution with contour resampling.

From the evolution of the contour in Fig. 4.2a to the final contour shown in Fig. 4.2b, we obtained the trend line for each termination criterion as shown from Fig. 4.3a to 4.3f. In this evolution process, visual inspection of the evolving contour suggested that an accurate and efficient termination metric should terminate the evolution between iteration 800 to 1000.

From the evolution of the contour in Fig. 4.2a to the final contour shown in Fig. 4.2c, the trend lines for image difference, group average and GAD are shown in Fig. 4.4a to 4.4c respectively.

Referring to Fig. 4.3, minimum distance and maximum distance (Fig. 4.3d and 4.3e) have shown to be non-decreasing metrics which are difficult to be used directly for automatic

termination. Although median distance (Fig. 4.3f) is a decreasing metric, it shows a great extent of fluctuation. Sum of norm, sum of distance and APM (Fig. 4.3a to 4.3c) show similar decreasing trends although their ranges are different. APM can be a representative of these three metrics because its valleys are nearest to zero. It is useful to highlight that these three metrics hit their first low around iteration 600. On one hand, if a threshold were to be set at the level of the metric at iteration 600, premature termination could occur (because accurate termination for this segmentation should occur between iteration 800 to 1000 based on observations). On the other hand, if a threshold were to be set at a very low value of a metric not covered in these plots, the time taken to termination could be unacceptably long.

The problem of pixel movement based termination criterion is not only its inability to be applied when contour resampling is present, but also its difficulty to find an accurate and efficient termination threshold. From our study, the value of the metric corresponding to the first lowest valley of a pixel movement based metric is highly dependent on the shape of the target structure. The unpredictability of pixel movement based metric is high with respect to target shape change. The problem of shape dependence of pixel movement based criterion will be evident when change in target shape causes the failure of the criterion to terminate (see following subsection: Time Accuracy Study).

In contrast to the pixel movement based termination metrics, although image difference is a non-decreasing metric (Fig. 4.4a), its trend is predictable (based on our observations: it tends to approach zero). Therefore, using image difference we are able to derive a

steadily decreasing metric, GAD. It is evident from Fig. 4.4c that a threshold value can easily be chosen by a value near to zero so that the evolution can terminate automatically at iteration 800 to 1000.

Fig. 4.3. Trend lines of pixel movement based metrics. (a) sum of norm, (b) sum of distance, (c) average distance (APM), (d) maximum distance, (e) minimum distance, and (f) median distance versus number of iterations.

Fig. 4.4. Trend lines of metrics related to Group Average Difference. (a) Image difference, (b) group average, and (c) group average difference, as defined in Eq. 4-10, 4-12 and 4-13 respectively.

Time-accuracy Study

This study examines the speed and accuracy of each termination method. The segmentation speed and accuracy of the methods were measured by the time of termination and DSC respectively. To obtain objective measures of the time of termination and DSC across different target shapes, the values of time and DSC were

averaged across different test shapes and only the average values are used for comparing different termination methods. The test images consist of six synthetic binary images having test shapes of varied complexity (Fig. 4.5). The images are of moderate size, 325×273 pixels.

Initial contours for all the synthetic images are all circles of radius 45 pixels centered at the center of the image. This study consists of two parts: the first part examines the termination methods when all associated parameters (including smoothness parameters of contour, time step, APM threshold, GAD threshold and group size) are optimized to the specific shape *and* the specific method. The aim of this part of the study is to compare APM with GAD in their best achievable performance. The second part of this study examines the feasibility of the methods for automatic implementation in which different shapes need to be segmented. This part of study compares APM with GAD while each parameter (including smoothness parameters of contour, time step, APM threshold, GAD threshold and group size) is set to a unified value for all shapes in an anatomical structure. In this study, the best combination of parameters is chosen for each method for comparison.

Fig. 4.5. Comparison of segmentation results using GAD and APM. Leftmost column: initial contours; middle column: final contours of APM terminated evolutions; rightmost column: final contours of GAD terminated evolutions. All final contours are from automatic implementation.

Results from the first part of the study are shown in Table 4.1. APM termination method gives accuracy, measured by DSC, of $97.55 \pm 1.59\%$. GAD termination gives accuracy of

97.55 $\pm$ 1.65% which is approximately the same as that of APM. The average time taken using GAD method to terminate is 69 seconds and the average time for termination using APM is 24 seconds. In this case, which all the parameters are optimized for the specific shape and specific method, APM method is 3 times faster than GAD method.

Table 4.1. Comparison of time taken, number of iterations to termination and accuracy for 6 synthetic test shapes using the conventional method and the proposed method. All parameters optimized for each shape and each method.

Conventional method (optimized)					Proposed method (optimized)				
APM					GAD				
	Threshold	Time (s)	Iteration	DSC	Threshold	Grp size	Time (s)	Iteration	DSC
Indented circle	0.0466	13	363	0.9890	0.25	15	22	390	0.9881
Heart	0.0533	15	436	0.9896	0.25	50	20	350	0.9904
Quadrilateral	0.0699	39	1110	0.9849	0.25	50	66	1200	0.9838
Triangle	0.0597	11	323	0.9717	0.25	100	187	3700	0.9736
Star	0.0470	16	479	0.9480	0.25	100	27	500	0.9457
Fish	0.0469	47	1361	0.9700	0.25	50	93	1550	0.9714
Mean		24	679	0.9755			69	1282	0.9755
St. dev.		15	442	0.0159			65	1281	0.0165

To test the time and accuracy of the termination methods considering its robustness with variation of target shape in automatic implementation, the second part of this *time-accuracy study* tests each method with fixed value of parameters across all test shapes. For APM method, the combination of parameters including smoothness parameters of contour, time step, and APM threshold that gives the best average accuracy of all shapes are chosen for comparison. Similarly for the GAD method, the combination of parameters including smoothness parameters of contour, time step, GAD threshold and group size that gives the best average accuracy of all shapes are chosen for comparison. Table 4.2 gives comparison results of APM with GAD in their best variation studied, while keeping parameters unified across all shapes.

Table 4.2. Comparison of time taken, number of iterations to termination and accuracy for 6 synthetic test shapes using the conventional method and the proposed method. All parameters unified for all shapes; best-accuracy combination of parameters for each method.

Conventional method (unified parameters)					Proposed method (unified parameters)				
APM					GAD				
	Threshold	Time (s)	Iteration	DSC	Threshold	Grp size	Time (s)	Iteration	DSC
Indented circle	0.0466	12	363	0.9891	0.25	100	39	700	0.9902
Heart	0.0466	29	850	0.9912	0.25	100	56	1000	0.9896
Quadrilateral	0.0466	336	10000	0.9858	0.25	100	105	1900	0.9847
Triangle	0.0466	338	10000	0.9781	0.25	100	101	2000	0.9742
Star	0.0466	335	10000	0.9521	0.25	100	268	5100	0.9536
Fish	0.0466	22	665	0.9689	0.25	100	41	700	0.9695
Mean		179	5313	0.9775			102	1900	0.9770
St. dev.		173	5137	0.0149			86	1670	0.0142

In this second part of *time-accuracy study*, the GAD method gives slightly lower accuracy ($97.70 \pm 1.42\%$) compared to the DSC accuracy of APM ($97.75 \pm 1.49\%$). However, the accuracy difference is neither large nor significant statistically (Fig. 4.5). The average time for termination using GAD is 102 seconds, almost half of the average time using APM, which is 179 seconds.

In this automatic implementation which all termination parameters were set to constant for all images, the results reveal the *shape sensitivity* nature of the APM termination method because in three out of six test images, the termination metric did not reach the threshold set before reaching the maximum number of iterations allowed. The GAD method is relatively shape insensitive because the contours are able to terminate before hitting the maximum number of iterations. Because of this shape sensitivity nature of pixel movement based termination metric, its average time for termination become longer compared to the GAD method in automatic implementation.

Evaluation of Termination Methods on Medical Images

For the purpose of testing the termination criterion in situations where anatomical structures with different shapes in different subjects or at different slice locations, three exemplar anatomical structures (sagittal femur, sagittal tibia and axial humerus) were segmented using sample anonymized medical images. The sizes of all medical images used in this section are the same, 320×320 pixels. All medical images were preprocessed to binary images before the contour evolutions. The preprocessing uses optimal threshold for each image for best delineation of the desired structure. The unified termination parameters (i.e. APM threshold, GAD threshold and group size) for automatic implementation were obtained by optimizing between the accuracy and the time for termination, for each termination method. APM threshold was optimized for each anatomical structure.

Because of the large extent in which the anatomical target shape can vary across different slices in a volumetric datasets, and that the purpose of this study is to compare termination methods but not initialization methods, the initial contours are not fixed across different target shapes, but are consistent across different termination methods for the same target shape. Being consistent in usage of initial contour ensures that the comparison of termination methods is unaffected by variation of initial contour. Initial contour for each medical image is defined using interpolated cubic spline curve by receiving a number of user specified points. Arbitrary number of points is used to define the initial contour for the medical images so that the initial contour for each image is fairly smooth and near to but not at the intended final location (where a contour should

converge). This way of initialization is to simulate the contour evolution at and only at the termination stage to observe termination pattern using APM and GAD termination criteria, respectively.

Similarly, for each shape in the medical images, smoothness parameters and time steps are optimized (in another study) for each target anatomical structure and used during contour evolution of medical images of that structure, consistently across different termination methods. These parameters are implemented in the settings for active contour runs in examining and comparing the conventional termination method (APM) and proposed method (GAD).

Table 4.3 shows the comparison of the automatic implementation for APM and GAD termination criteria. Both methods of termination achieved very good accuracy measured by DSC. GAD has an average accuracy of $96.77 \pm 0.92\%$ and APM has an average accuracy of $96.75 \pm 0.94\%$. There is no significant difference in accuracy using these two methods. However, the shape sensitivity of APM become apparent when 6 out of 9 images were unable to terminate before reaching the limit for number of iterations; in contrast, in the case of GAD termination criterion, none of the images fails to terminate before hitting the iteration limit. Shape sensitivity indicates a need for shape-specific tuning of the termination metric. Without shape-specific tuning, a shape-sensitive termination method can result in long segmentation time (which is not the optimal segmentation time the method can possibly achieve). In this automatic implementation, the total segmentation time of all images tested using APM is 2121 seconds, whereas the

total time using GAD is 607 seconds. In this case, there is about 60% total time reduction using GAD compared to using APM. The average segmentation time of all images tested using APM is 236 ± 217 seconds, whereas the average time using GAD is 67 ± 50 seconds. Both total segmentation time and average segmentation time of GAD represent about 60% time reduction compared to APM. The coefficient of variation (COV) of the segmentation time is 92% using APM and 75% using GAD. The COV can be regarded as a measure of relative shape sensitivity for comparing the methods because the segmentation time would have a large variation from its mean if the termination criterion cannot terminate the contour within the maximum allowed number of iterations in some cases. The larger COV of segmentation time resulted from using APM suggests that APM is more shape sensitive than GAD.

Table 4.3. Comparison of time taken, number of iterations to termination and accuracy for 9 medical images of three anatomical locations using the conventional method and the proposed method in automatic implementation.

Conventional method (unified threshold)					Proposed method				
APM					GAD				
	Threshold	Time (s)	Iteration	DSC	Threshold	Grp size	Time (s)	Iteration	DSC
Femur lateral	0.0468	418	10000	0.9731	0.25	50	76	1000	0.9710
Femur middle	0.0468	496	10000	0.9655	0.25	50	42	500	0.9660
Femur medial	0.0468	6	194	0.9733	0.25	50	14	200	0.9730
Tibia lateral	0.0431	187	10000	0.9802	0.25	50	133	2500	0.9805
Tibia middle	0.0431	544	10000	0.9742	0.25	50	155	1700	0.9741
Tibia medial	0.0431	6	270	0.9726	0.25	50	61	1000	0.9746
Humerus proximal	0.0460	332	10000	0.9533	0.25	50	40	600	0.9536
Humerus intermediate	0.0460	13	345	0.9584	0.25	50	77	1100	0.9580
Humerus distal	0.0460	119	10000	0.9567	0.25	50	9	200	0.9586
Total		2121					607		
Mean		236	6757	0.9675			67	978	0.9677
St. dev.		217	4865	0.0094			50	745	0.0092

*middle refers to the location of intercondylar slice in knee MRI

The final contours for four representative results using real medical images are shown in Fig. 4.6. Their corresponding trend lines for both APM and GAD evolutions are shown in

Fig. 4.7. There is no visible difference on the image between the segmentation results terminated by APM and GAD.

The tests of GAD termination criterion using medical images ascertain the decreasing trend of the proposed metric (Fig. 4.7). These tests further illustrate the difficulty in selecting a specific APM threshold because APM threshold is highly shape sensitive, i.e. requiring different threshold to terminate optimally. For example, the optimum APM termination threshold used for femur is optimal only for medial slice of femur. For the middle slice of femur which the target shape is different (not to confuse medial with middle), the threshold causes failure in terminating within maximum number of iterations. On the other hand, using a threshold optimum for the middle slice for the medial slice may in turn result in premature termination in the medial slice.

In contrast, without laborious manual search for a specific threshold for specific shape, an evolution is able to terminate within reasonable number of iterations (in this study we consider 10 000 iterations as maximum for reasonable number of iterations) for all nine shapes tested using GAD (Table 4.3).

Fig. 4.6. Results of 3 representative segmentations of anatomical structures. Leftmost column: initial contours; middle column: final contours of APM terminated evolutions; rightmost column: final contours of GAD terminated evolutions. Top row: segmentation of femur middle slice using sagittal knee MRI; 2nd row: segmentation of femur medial slice from the same dataset; 3rd row: segmentation of humerus intermediate slice from axial shoulder MRI; bottom rows: segmentation of humerus distal slice from the same dataset.

Fig. 4.7. Comparison of trend lines using APM and GAD. Left column: trend lines for APM (fixed threshold); right column: trend lines for GAD (unified group size and threshold). Top row: trend lines for segmentation of femur middle slice; 2nd row: trend lines for segmentation of femur medial slice; 3rd row: trend lines for segmentation of humerus intermediate slice; bottom row: trend lines for segmentation of humerus distal slice. The coordinates of first valleys of APM lines are labeled for over-evolved cases.

4.1.3. Discussions and Conclusion

Segmentation Time

Significant variation in the shape of a target structure can be encountered in practical situations such as: 1) same anatomical locations of different patients; 2) different pathological severity at the same anatomical location of the same or different patient(s) 3) different slices of the same volumetric dataset. Hence in an automatic implementation to segment an anatomical structure, it is important that the termination criterion in use is insensitive to this shape variation and terminate the snake evolution in a reasonable time and accuracy.

Because the time for termination of the proposed method is longer than that of optimized-threshold APM termination (Table 4.1), it is probable that one favors the use of optimized-threshold APM to terminate a contour. However, in practical application of automatic segmentation task which involves extracting structures of different shapes, it is unlikely that the user would search for the optimized parameters for each shape because it involves the work as laborious as manually segmenting the target structure. In automatic segmentation task, due to variation of shapes across different images, a termination method that does not need user intervention in searching for the correct threshold value is usually required. Table 4.2 provides the results which the methods run without user intervention across different shapes. The results suggest that due to the shape sensitive

nature of APM termination, its segmentation speed become disadvantageous in automatic implementation. Although GAD has one more parameter involved (the group size), its relative insensitivity to change of target shape suggests that the tuning required is considerably less than that of the application of APM, which only has one parameter to tune but very sensitive to change of target shape. We have shown that the average time for termination using GAD in automatic implementation can be as short as half that using APM in the synthetic images. Tests in medical images ascertain the notion of shape sensitivity of APM and that its time for automatic termination is longer than that using GAD.

Contour Resampling

All tests of APM termination criterion in this study were carried out without contour resampling at every time step because contour resampling will cause the loss of vertex correspondence between the contours of subsequent time steps. This is an inherent restriction of using pixel movement based termination criterion. On the other hand, for stable evolution without redundant and/or insufficient vertices to characterize the contour, all tests of GAD termination criterion were carried out with contour resampling at every time step. Since the GAD termination criterion can work with contour resampling and APM criterion cannot, the pixel movement based termination criteria such as APM are considered to be sensitive to the use of contour resampling.

Number of Iterations

The number of iterations for a parametric active contour depends on several factors such as the image size, the object size, use of normal force selection, use of contour resampling, and the shape of the object relative to the initial contour. In this study, the large numbers of iteration are mainly due to the *normal force selection* formulation that keep the contour fluctuate around the exact boundary sought; and in medical images, partially caused by the relatively large size of images. Normal force selection is a formulation important to avoid vertex clustering and vertex dispersion while evolving the contour. The formulation also provides a more predictable movement pattern for evolving snake since only unit normal force is applied onto the contour at any time. Because the contour is formulated using normal force selection, the movement of contour will not decrease to zero but will settle at pseudo equilibrium state which the contour fluctuate around the exact solution within approximately ± 3 pixels. While gaining the advantage that the contour vertices do not cluster/disperse because they do not move in the direction tangential to the contour, the associated effect which may be disadvantageous is the large number of iterations before the contour can be terminated by any criterion set objectively.

Image Size

Relatively small image is used for *trend line study* because use of a small image will result in the contour reaching equilibrium position at smaller number of iterations. Therefore, a relatively low limit of iteration (1200) is sufficient to facilitate analysis of the entire trend of termination metric. For *time-accuracy study*, this study examined the methods of interest with images of moderate size. The limitation (in terms of time and

accuracy) of the proposed metric will be evident on small (or thin) structure on large image, or any structures on small image. For these instances, the pixel movement based metrics could perform much better in speed (including segmentation speed and evolution speed) and accuracy compared to GAD although their problems of shape sensitivity and sensitivity to contour resampling remain.

A Similar Termination Metric

A metric similar in concept to our proposed metric based on the change of contour length was recently proposed by Wang et al. [12]. Our metric is based on area difference of the contour. We use the group size g to remove the fluctuation of the base metric (image difference) because the fluctuation of it is large and the number of iterations involved in parametric active contour is large; Wang et al. use a threshold of iteration, $T_{iteration}$ similar in concept to our group size. Since the number of iterations to the target shape for their geodesic active contour is small and the fluctuation of their metric is also small, they do not use $T_{iteration}$ to average the base metric but use it as the number of iterations that if the base metric (change in contour length) is below threshold for $T_{iteration}$ iterations, termination occurs. Both our proposed GAD and Wang's termination criterion are insensitive to the type of contour.

Limitations of Study

This study examines the proposed termination criterion from the point of view of application to GVF active contour with normal force selection. Our focus in this study is

primarily to illustrate the advantages of the proposed metric compared to pixel movement based metrics especially in situation where contour resampling is necessary.

The proposed method itself is applicable to all kinds of active contour regardless of its force balance formulation, but the group size is a parameter needs to be adjusted when the formulation or numerical implementation of active contour is changed because the extent of fluctuation of the contour at desired boundary is changed. Although the proposed metric shows good potential to be applied to other variations of active contours (e.g. geodesic active contours [80], the balloons [81] and region-based active contours [169]), its comparative performance (compared to other metrics) in those applications remain to be studied. This study only compared the termination methods using 2D active contour but not active surface in 3D. An extension of the proposed method would be to measure GAD based on volume difference (2D is based on area difference).

All images used in this study possess defined boundaries. We compared the termination criteria of interest in cases where defined boundaries are present because even with defined boundary, the contour can fluctuate when it reaches the boundary, as shown in the results. Consequently, terminating the evolving contour automatically is not trivial. The study does not intend to address the challenge in accurately terminating an evolving contour in the cases of ambiguous boundary which is the subject of active contour formulation (e.g. introduction of a region-based term in the energy functional).

This section has presented a new termination criterion called GAD designed for parametric active contour. We assessed the robustness of the criterion by considering its trend, speed, accuracy, shape insensitivity and insensitivity to contour resampling in comparison to a pixel movement based criterion. The proposed termination criterion showed high predictability of decreasing trend in a typical evolving active contour. In automatic applications, it has considerably higher average speed to termination than pixel movement based termination criteria with comparable accuracy. Given that a fixed threshold of termination metric was used across different target shapes, the failure (to terminate) rate of the proposed termination criterion is very low compared to highly shape sensitive pixel movement based termination metrics. We have also demonstrated the insensitivity of the proposed termination criterion to contour resampling. Since the proposed criterion only depends on the segmentation outcomes of the contour but not the formulation of the contour, it should be applicable to different types of active contour/surfaces.

4.2. Developing Parametric Active Contour with Centroid Force⁵

MR image is useful for diagnosis and treatment evaluation of the cartilage for patients suffering cartilage disease. Trend in quantitative imaging and increasing amount of data being studied have called for the need to automate the diagnosis in the near future [40]. Analysis of cancellous bone after segmentation will further our understanding of the

⁵ Part of this section has been published in : T.K. Chuah, C.W. Lim, C.L. Poh, and K. Sheah. Bone segmentation of magnetic resonance images by Gradient Vector Flow active contour with atlas based centroid forces. 2010. DICTA, Sydney, NSW. Permission of reprint granted in the copyright agreement.

change of cancellous bone to progression of bone related diseases [170-171]. This includes the extraction of cancellous bone of the femur (the region of thigh bone where bone marrow lies, we denote this part as *cancellous femur* hereon) for characterization of its material property for prognosis of osteoarthritis and osteoporosis, such as examining the texture of bone marrow lesion. This section of thesis presents an effective way to initialize active contour for segmenting the cancellous femur to facilitate analysis of bone marrow and potentially localizing other structures including cartilage.

A major difficulty in developing automatic bone segmentation is the intensity ‘leakage’ near the cartilage boundary, especially at the anterior site of femur. From our experimentation, method using Otsu’s thresholding [167] followed by seeded region growing has proven to be ineffective for the purpose of automatic cancellous bone segmentation because of intensity overlap with surrounding tissue. This inability of region growing method is mainly because the targeted region tends to grow through narrow sites of intensity overlap that appear to join the targeted region with undesired region. These sites of intensity overlap generally have ambiguous boundaries.

The delineation of ambiguous boundary in the segmentation of cancellous femur remains a challenge. Ambiguous boundary can be determined from atlas by registration [172], or by constraints imposed on a contour [83]. Pure atlas registration segmentation techniques [173] either require long processing time (for non-rigid registration), or that they are unable to locate the exact image boundary at some apparent locations. Parametric active contour (AC) is robust at delineating all apparent boundaries with a closed contour, and it

segments one object at time. However, one problem with AC is that sites of ambiguous boundary are usually erroneously delineated because edge information is either incomplete or missing. Improvements of active contour in attempt to address the problem include balloon [81] and region-based active contour [174].

Another problem with AC is that formulation of parametric AC does not suggest any fast and reliable way to initialize the contour or to push the contour to the regions of external force coverage. One of the gaps between current developments of AC to automatic segmentation lies in the inability to initialize the contour automatically [175]. To address the problems of ambiguous boundary and initialization in cancellous bone segmentation, this section presents an atlas based parametric AC for automatic segmentation of proton density (PD) weighted knee MR images. An atlas can be defined as a model generated from multiple subjects; or one representative model chosen from a database, that most resembles the subject of interest [173, 176]. For simplicity, atlas here refers to a single representative model derived from one subject. PD weighted images provide excellent contrast among the various tissues of the knee joint. The proposed technique utilizes an atlas to constraint the contour at regions where edge based forces are missing, as opposed to setting a mathematical formulation to constraint the contour which may be heavily reliant on image primitives (e.g. edge-based or region-based formulation) or extensive experimentation (e. g. tuning the balloon force). An atlas based force which we called centroid force is added as an external force in our proposed technique.

4.2.1. Methods

Image Processing Procedures

A 3D atlas of the femoral cancellous bone was constructed by manually segmenting a proton density weighted MR knee dataset. The atlas is a binary image in 3D containing the shape information of the femoral cancellous bone. For simplicity, a single atlas is used in this study [173]. In the atlas, the femoral cancellous bone is marked by binary 1 and background by binary 0. This atlas was separately registered, by user interaction in *3D-Doctor* [177], to each 3D dataset of the subject. This interactive registration involves the adjustment of uniform scaling and rigid transformation (scalable rigid) parameters to achieve maximum spatial overlap between the atlas and the cancellous bone of the subject image. Because the focus of this study is on the investigation of whether the atlas based centroid force would address the problems with ambiguous boundary and initialization, the registration procedure is done interactively instead of automatically which may require finding model-to-image (i.e. atlas to image) point correspondence for registration [178]. The atlas registered to the space in subject image serves as a constraint to the AC at all ambiguous boundaries and provides localization information for initialization of the contour.

Subject images registered with atlas in 3D are then segmented on a slice-by-slice basis. Slices at the intercondylar location often show intensity overlap with cartilage at the anterior region, and with the ligament insertions at the posterior region. Hence, current study analyzes the slices at the intercondylar position of the femur. Gradient Vector Flow

(GVF) active contour is used for locating the boundaries of the femoral cancellous bone [82]. Fig. 4.8 shows an illustration of GVF force field. GVF is chosen because it is robust in detecting edge and providing a wide capture range. Capture range is the distance which the GVF or other force field can reach; a wide capture range means the initial contour can be placed anywhere in the vicinity of the desired edge; a narrow capture range means the initial contour needs to be placed very near to the desired contour. For obtaining external force of GVF, the image to be segmented (target image) is first binarized using Otsu's threshold [167]. By carrying out binarization, edge information can be extracted out and erroneous edge information (caused by Rician noise in MRI image, and by bone marrow which is not a structure of interest. Note that the noise and the undesired structure cannot be removed effectively using conventional way of Gaussian filter in active contour framework) resulted from using grey-level gradient directly can be minimized. Furthermore, the noise in the binary image is filtered using connected component labeling and area expunction. Connected components of area smaller than that of the size at 95 percentile are removed. This procedure removes undesired small patches at random locations. To extend the area of image gradient, the filtered binary image is then subjected to Gaussian smoothing of a 5×5 kernel with $\sigma = 2.0$. This Gaussian filtered image is then used for obtaining the gradient force, i.e. the external force of snake formulation by Kass [75]. Only the directions of gradient forces are the used to initialize a force-field diffusion process in GVF. The diffusion outcome is controlled by the GVF time step, ζ_{GVF} .

Fig. 4.8. Illustration of image gradient and GVF field. (a) An example of grey level image. (b) binarized image of the grey level image. (c) Image gradient of (b). (d) The GVF field extended from (c) which covers more area in the image.

A modification to the GVF external force is done by adding centroid force. This centroid force is used to pad the area where the GVF external forces are missing or too small. Since the atlas is registered to the subject image, each 2D subject image has a corresponding atlas slice. The registered atlas can barely be used directly as the segmentation result because many of the boundary points of the atlas still do not intersect the boundary points in the structure of interest. Center of mass (centroid) of an atlas slice is calculated [179]. This centroid and the atlas slice are utilized in two ways:

1) *Centroid Force*: an external force field originated from the centroid. The field contains the forces that point outwards when they are within the foreground of the atlas, and inwards when they are within the background of the atlas. The centroid force $\mathbf{F}_c$ can be described by the equations:

$$\mathbf{F}_c(x, y) = m(x, y)\mathbf{n}(x, y) \quad (4-14)$$

$m(x, y)$ above is a Gaussian shape function

$$m(x, y) = \exp(-r^2 / \sigma_c^2), \quad (4-15)$$

Where σ_c defines the deviation of the Gaussian shape; $\mathbf{n}(x, y)$ is a unit normal vector pointing towards or away from the centroid; x, y are pixel coordinates; r is the distance of point (x, y) from the centroid. Note that $m(x, y)$ can also be chosen as the distance transform of the binarized image.

Our formulation consists of a combination of GVF force and centroid force. The complete external force field can be described by the following equation:

$$F_{complete}(x, y) = \begin{cases} \mathbf{F}_{GVF}(x, y), & \text{iff } |\mathbf{F}_{GVF}(x, y)| \geq T \\ \mathbf{F}_c(x, y), & \text{iff } |\mathbf{F}_{GVF}(x, y)| < T \end{cases} \quad (4-16)$$

T is a threshold value for the magnitude of GVF force used to determine when centroid force will be used instead of GVF force. Because GVF forces of small magnitude usually do not give correct direction for contour evolution, centroid force will be used instead. Hence, T is usually chosen to be a small value. In our application, T is arbitrarily set to

0.01. If T is chosen to be 0, by which one has a continuous mixing of two forces, the evolution will be slower due to presence of small GVF forces.

2) *Contour Initialization*: Initialization of contour is done automatically by placing a circle centered at the centroid with a radius of $0.5R_{min}$. R_{min} is the minimum radius from the centroid to the edge of foreground of the atlas. Fig. 4.9 shows an illustration of centroid force when the atlas is a circular shape.

Fig. 4.9. Centroid force (arrows) for a filled circle with radius of 5 pixels on a 32×32 image (shown here as translucent mask).

After padding the GVF field with the centroid force field, the contour is evolved to locate the femoral cancellous bone. Contour evolution is regulated by using only unit force at the direction normal to the contour. Contour resampling is done iteratively every time step of evolution as suggested by Lobregt and Viergever [163]. The contour was then automatically terminated based on a criterion set for area bounded by the contour in subsequent iterations, which was denoted as GAD in previous section [180].

Images and Analysis

A total of twenty one slices located at the intercondylar location from seven 3D datasets of sagittal knee proton density weighted MR images were selected for studying the proposed method for bone segmentation. This location is selected because the slices here are most susceptible to the leakage problem in segmentation as compared to slices at medial or lateral condyles, due to the presence of thin cartilage, and insertions of ligaments. The images are of size 512×512 , in-plane resolution of 0.27 mm and slice thickness 1.5 mm. The twenty-one slices were considered separately during analysis. The accuracy of the segmentation method proposed was compared with expert manual segmentation (the ground truth) using sensitivity, specificity and DSC.

4.2.2. Results

Fig. 4.10 shows how the atlas based centroid force completes the missing force field of GVF. In Fig. 4.10b, top left corner and bottom right corner show the GVF field (black arrows) that enables the contour to delineate the sought boundaries accurately; top right corner to bottom left corner of the figure show no GVF force originally, and are now padded with atlas based centroid force (blue and red arrows). This centroid force field completes the missing boundary information which cannot be obtained using only GVF formulation. Consequently, the contour was able to settle at the approximated boundary given by the atlas. Fig. 4.10c shows the segmentation results at the proximal part of femur, with and without centroid force. Notice the difference in accuracy at the anterior tip and at the ligament attachment.

Another comparison of conventional GVF and atlas based GVF is shown in Fig. 4.11. Top row of the figure shows the initial contour and the final contour using conventional GVF active contour. The final contour is smaller than the initial contour and does not change shape because external force (the GVF field) cannot be detected by the contour (capture range of the external force is small). If we initialize with a larger circle, as shown in the second row of Fig. 4.11, some of the external forces can be sensed by the contour, but the final contour still cannot effectively detect the external field at the proximal and distal (upper and lower) end of the femur because they are too far away and there are missing force fields between the initial contour and these ends. Bottom row of Fig. 4.11 shows the initial and final contour of an atlas based AC we proposed. Even though the initial contour of this evolution (Fig. 4.11e) is smaller than that of Fig. 4.11a (i.e. more unlikely to be within the coverage of GVF field), the contour can move toward the distal and proximal ends of the femur because of the atlas based centroid force. The initial contour in Fig. 4.11e is chosen to be smaller than Fig. 4.11c to show that the initial contour does not need to be large enough to be within the GVF field.

Fig. 4.10. Combination of centroid force and GVF force. (a) Left: Ambiguous boundary of the knee MR image highlighted by a zoom-in box; right: the Otsu's thresholding result of the image showing 'leakage' of the bone at the highlighted site. (b) External Force field of GVF (black) modified by centroid force (blue and red, blue indicates inward pointing forces, red indicates outward pointing ones), at the ambiguous boundary of the zoom-in box in (a). Transition between red and blue is the estimated boundary. (c) Image at the left shows the proximal part of femur with the initial contour; image in the middle shows the segmentation by GVF field without centroid force; image at the right shows the segmentation by the proposed method which combines centroid force and GVF force.

Fig. 4.11. Segmentations with and without centroid force. The first column shows initial contours, second column shows final contour terminated based on area bounded by the contour. First row is the segmentation by GVF without centroid force, initialized with a small circle. Second row is the segmentation by GVF without centroid force initialized with a larger circle. Third row is the segmentation of GVF with centroid force initialized with the same small circle.

The accuracy measures of femoral cancellous bone from the 21 slices of interest are summarized here: the sensitivity is $97.4 \pm 1.9\%$; specificity is $99.6 \pm 0.1\%$; DSC is $96.7 \pm 1.1\%$.

4.2.3. Discussions and Conclusion

GVF active contour is known for its robustness in tracking inward concavity and the extended capture range compared to the basic formulation of external force based solely on the edge map. The edge map $f(x,y)$ used in this study is defined as follows:

$$f(x,y) = |\nabla [G_\sigma(x,y) * I(x,y)]|^2 \quad (4-17)$$

Where $I(x,y)$ is the filtered binary image of the MR image; G_σ is the Gaussian smoothing kernel with standard deviation σ ; $*$ is the convolution operator and ∇ is the gradient operator. However, in the filtered binary image $I(x,y)$ of interest in this study, the large image size (512×512) makes GVF insufficient to cover the entire structure of interest. For cases which the GVF field covers the whole structure of interest, a semi-automatic initialization as suggested by [175] can be employed. For cases which the GVF is insufficient to cover the whole structure of interest, we propose this atlas based centroid force method. This method fills up the force field where GVF force is missing and facilitates both automatic initialization and constraining of contour at ambiguous boundaries. This constraint imposed using atlas based force can avoid the situation that the contour is over reliant on fine-tuning of the contour smoothness parameter.

Over reliance on tuning of contour smoothness parameter is a hurdle to automatic process because the operator has to continuously look for different smoothness values for different images. Furthermore, one set of smoothness parameter values does not necessarily ensure accurate segmentation at all locations because some locations of ambiguous boundary maybe smoother than the other. Another possible solution to missing external force field for AC is the balloon force [81] that can inflate the contour to achieve the same effect as Fig. 4.11f. However, balloon force cannot accurately approximate the ambiguous boundary like the atlas force unless tuning of the magnitude of inflation force is strenuously done. In contrast, atlas dependent force does require tuning of force magnitude (for contour evolution) to achieve approximation of the ambiguous boundary.

Another comparable method is the region-based strategy of active contour suggested by Ronfard [174]. The snake evolution of Ronfard's method evolves like a balloon but is more comprehensive than a balloon [81] because the evolution in homogeneous region is controlled by region (both inside and outside) information. The region-based snake solves the problem of limited capture range for sharp concavities but it does not readily recognize ramped edges. Our proposed method, however, incorporates prior information (atlas information) for recognition of ramp edges and image information for recognition of step edges. Due to the lack of prior information in the region-based snake, automatic initialization is not possible in [174].

The proposed method was intended to overcome hurdle in automated segmentation process. Therefore the need to initialize the contour slice by slice is taken over by the automatic process of defining a circle centered at the centroid with a radius obtained from the atlas. The problem of insufficient capture range and initialization of contour are solved simultaneously by this method because the initial contour, regardless of how small its radius, will always approach the meaningful GVF field directed by centroid force. This robustness of the method in locating GVF field can be visualized by looking at the external force diagram shown in Fig. 4.9, in which as long as an initial circular contour is placed within the target region and centered at the centroid, the centroid force will push the contour outward. Furthermore, although initial contour in Fig. 4.11e is smaller than that in Fig. 4.11a, the final contour can still detect the GVF field as shown by the final contour in Fig. 4.11f; unlike in the absence of centroid field, the larger contour (Fig. 4.11c) failed to detect the GVF field (Fig. 4.11d).

GVF without centroid force requires careful manual initialization of contour in order to avoid segmentation as unsatisfactory as that shown in Fig. 4.11b; whereas using GVF with centroid force, automatic initialization is applicable and gives satisfactory results. It is usually not practical to segment the target structure using GVF only especially for image of considerable size (e.g. larger than 256×256) because the final contour depends on how large the ambiguous boundary is, and how accurate the manual initialization is.

In this study, manual registration was performed. Hence, automatic registration will be applied in the future to achieve fully automatic segmentation. The 2D AC and centroid

force used in this study can also be extended to 3D. A larger population will also be used to form the atlas.

This section has presented a practical method to effectively localize the capture range of GVF active contour in relatively large image using atlas based centroid force for active contour after scalable rigid registration. The atlas is used in this study to constraint the contour at regions where edge based GVF forces are missing. Ambiguous boundaries at the interface between cancellous bone and its exterior were approximated by the shape information of centroid force derived from atlas. The use of atlas in initialization and constraint of AC provides a potential way towards automatic segmentation.

Chapter Summary

This chapter has presented two contributions to the application of parametric active contour automatically to the segmentation of bone in knee MRI. The first is the proposed termination criterion, GAD, which works more efficiently with almost the same accuracy as the conventional pixel movement based criterion. The second is the initialization strategy proposed to go around various challenges such as: requirement of a model constructed from many subjects, and/or leakage of active contour to ambiguous boundaries. The strategy offers an alternative to initialize and constrain evolving active contour for MRI. The two contributions have provided improvements to the use of active contour towards an automated segmentation system in CAD. Nevertheless, more studies

and validations will need to be done before the segmentation resulting from the proposed strategies can be applied in clinical routine.

Chapter 5

Summary and Significance of This Thesis

Computer-aided diagnosis (CAD) is becoming increasingly important for medical practitioners to cope with the increasing amount and complexity of data as technology advances and the population grows. It can improve the efficiency of diagnosis and decrease the variability of diagnostic decision made. Quantitative analysis that requires image segmentation is an integral part of a CAD system.

This thesis has addressed the problem of damaged cartilage visualization. Cartilage is a thin structure and diagnosis of cartilage lesion and potential thinning can be difficult. CAD visualization system can provide detection of potential damage and a way for radiologist to quickly examine the potential damage. The thesis has presented a framework developed to visualize potentially damaged cartilage in 3D; and presented a study to investigate the relationship between percentage of cartilage damage (volume) to a histogram-based parameter.

Besides, this thesis has also investigated the use of textural parameters in analyzing the presence of BML which is the one of the most common pathologies seen on the bone marrow. The thesis has shown that there are a considerable number of textural parameters that can potentially be used to aid medical decision making either directly as supplementing information, or be incorporated into a classification system that provides second opinion or as a screening tool. Following it, the possibility and feasibility of using textural information as second opinion has been studied. It was shown that the accuracy of using textural information in classifying whether a subject has edema can be highly accurate, provided that the size of edema is within the capture range of the system.

For advancing the automation of the segmentation task, this thesis applied a method that is efficient in the usage of data and computational resources: which is the use of single model as reference to automatically segment the bones. This thesis contributed to the advancement by developing termination criterion of active contour and proposed a way to effectively initialize and constraint active contour for the application of bone segmentation. The termination criterion developed has been shown to perform more efficiently while using it as a part of automated system.

In sum, the thesis has contributed to the development of analysis of bone marrow, segmentation of bone, and the visualization of cartilage. These contributions will help radiologists to be more efficient and accurate in their diagnosis, and aid them to focus on images that need more medical attention. More efficient diagnostic cycle will translate to

better patient care and lower medical cost for the general public. Any companies developing a CAD system could gain some insights from the results of investigations in this thesis. These will benefit the society as a whole.

Due to limitation of resources within in the course of producing this thesis, it was not possible to address each and every issue leading to a complete CAD system for musculoskeletal diseases. There is still room for advancement to reach the target of routine implementation of CAD the system for diseases such as knee OA. To implement the developed active contour segmentation system, tests and validation will have to be carried out in 3D with a larger sample size. The manual registration step that is required in the initialization of the Active Contour with Centroid Force will have to be automated to complete an automated segmentation system. The registration step may be automated in the future, for example, by registration of Gabor filter outputs. Nevertheless, the advancement made to terminate the active contour and in utilizing centroid force can widely be applicable to any kinds of image segmentation tasks.

A schema of complete CAD system for musculoskeletal diseases is show in Fig. 5.1. The figure shows how each component connects with the others. The system takes its input from the images and based its intelligence on statistics, then provides radiologists with useful and objective information such as location of potential disorder and the diseases class of a subject. Radiologist can then make report and decision based on this information along with his experience. The first component of the CAD system, which is the automatic segmentation system, is a widely studied area. Semi-automatic

segmentation is commonly studied and is an essential part of approaching the goal of accurate automatic segmentation. This thesis worked on the advancement of automatic segmentation system by studying termination, initialization and restriction of evolution of active contour. Moreover, the study of cartilage characterization developed the area in visualization of the structures and also intensity analysis of the cartilage. The analysed intensity, along with the texture of bones studied, could be put into a multivariate classification system to determine disease state and/or disease type.

Fig. 5.1. A complete schema for a fully automatic CAD system.

The classification system demonstrated for detection of BML could potentially improve the accuracy and decrease the work load of radiologists. A study that compares the performance of radiologists with and without the aid of the system would be the necessary next step to justify the use of the developed classification system. Furthermore, a comparison can be done to test the value of the visualization framework developed for viewing potential cartilage lesion, and also the use of supporting parameters such as the

histogram based parameters of the cartilage. In addition to radiologist's observation, other data such as patient scores, radiographs, arthroscopic findings and blood biochemical data could be used in future studies to further validate the results of the system.

One of the limitations of the study is that only proton density images were used for developing the system. For a system which can potentially provide classification of osteoarthritis from metastatic lesions, storage disease, metabolic disturbance and bone bruises; a system using more types of MR sequences (such as T1-weighted, T2-weighted, STIR and relaxation map) may be advantageous. Another limitation is that when studying texture of BML, we assumed that BML is a proxy to osteoarthritis which is not a proven hypothesis. Clinical and statistical study will have to be carried out to prove or disprove the validity of this assumption. If it is proven, then we can use texture of BML to diagnose osteoarthritis. The method presented in this thesis still only provides a way to detect BML, without the specificity about osteoarthritis, despite the fact that the data of the affected group were from patients who have osteoarthritis. The thesis only covered cartilage analysis, texture analysis and bone segmentation of the femur, leaving out tibia and patella. It is important to note that they are equally important in actual diagnosis. We focused on only femur because of the need to focus on single structure for initial development of methods for analysis. Furthermore, the thesis proposed a way to restrict the evolution of active contour using centroid force but tested on a very limited number of samples. Implementation of the method on a larger sample will be needed to demonstrate its robustness across different patients and acquisition centers.

For achieving a complete CAD system, modules for other structures such as ligaments, muscles, menisci are also of importance. With intensive research and enthusiastic funding in the field of CAD, the components developed will gradually be incorporated into systems used in actual medical settings to supplement human's diagnostic ability. CAD will then provide increased ability to display and/or interpret invisible image cue and interpret multidimensional information.

References

1. J.A. Buckwalter, C. Saltzman, and T. Brown, *The impact of osteoarthritis: Implications for research*. Clinical Orthopaedics and Related Research, 2004(427 SUPPL.).
2. N. Arden and M.C. Nevitt, *Osteoarthritis: Epidemiology*. Best Practice and Research: Clinical Rheumatology, 2006. **20**(1): p. 3-25.
3. T.P. Andriacchi, A. Mündermann, R.L. Smith, E.J. Alexander, C.O. Dyrby, and S. Koo, *A framework for the in vivo pathomechanics of osteoarthritis at the knee*. Annals of Biomedical Engineering, 2004. **32**(3): p. 447-457.
4. P. Conaghan, *Is MRI useful in osteoarthritis?* Best Practice and Research: Clinical Rheumatology, 2006. **20**(1): p. 57-68.
5. R.A. Stockwell, *Cartilage failure in osteoarthritis: Relevance of normal structure and function. A review*. Clinical Anatomy, 1991. **4**(3): p. 161-191.
6. P.G. Conaghan, D. Felson, G. Gold, S. Lohmander, S. Totterman, and R. Altman, *MRI and non-cartilaginous structures in knee osteoarthritis*. Osteoarthritis and Cartilage, 2006. **14**(SUPPL. 1): p. 87-94.
7. S.L. Hughes and D. Dunlop, *The prevalence and impact of arthritis in older persons*. Arthritis Care and Research, 1995. **8**(4): p. 257-264.
8. C.A. Marra, J. Cibere, R.T. Tsuyuki, J.A. Soon, J.M. Esdaile, L. Gastonguay, B. Oteng, P. Embley, L. Colley, G. Enenajor, and R. Kok, *Improving osteoarthritis*

- detection in the community: Pharmacist identification of new, diagnostically confirmed osteoarthritis. Arthritis Care and Research, 2007. 57(7): p. 1238-1244.*
9. *Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation --- United States, 2007--2009. MMWR Morb Mortal Wkly Rep, 59(39);1261-1265 2010 [cited 2012 10 Jul 2012]; Available from: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5939a1.htm?s_cid=mm5939a1_w.*
 10. *National Health Interview Survey, United States. 2007--2009. Publicly available from: <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5939a1.htm>.*
 11. H.P. Phua, A.V.L. Chua, S. Ma, D. Heng, and S.K. Chew, *Singapore's burden of disease and injury 2004. Singapore Med J, 2009. 50(5): p. 468-478.*
 12. S. Ministry of Health. *MOH Clinical Practice Guidelines 4/7. 2007 [cited 2009 19 June 2009]; Available from: http://www.moh.gov.sg/mohcorp/uploadedFiles/Publications/Guidelines/Clinical_Practice_Guidelines/OA%20Knees%20booklet.pdf.*
 13. P. Kiviranta, J. Töyräs, M.T. Nieminen, M.S. Laasanen, S. Saarakkala, H.J. Nieminen, M.J. Nissi, and J.S. Jurvelin, *Comparison of novel clinically applicable methodology for sensitive diagnostics of cartilage degeneration. European Cells and Materials, 2007. 13: p. 46-55.*
 14. D.T. Felson, *Developments in the clinical understanding of osteoarthritis. Arthritis Research and Therapy, 2009. 11(1).*
 15. E.B. Dam, M. Loog, C. Christiansen, I. Byrjalsen, J. Folkesson, M. Nielsen, A.A. Qazi, P.C. Pettersen, P. Garnero, and M.A. Karsdal, *Identification of progressors*

- in osteoarthritis by combining biochemical and MRI-based markers. Arthritis Research and Therapy*, 2009. **11**(4).
16. K. Doi, *Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. Computerized Medical Imaging and Graphics*, 2007. **31**(4-5): p. 198-211.
 17. D.M. Renz, J. Böttcher, F. Diekmann, A. Poellinger, M.H. Maurer, A. Pfeil, F. Streitparth, F. Collettini, U. Bick, B. Hamm, and E.M. Fallenberg, *Detection and classification of contrast-enhancing masses by a fully automatic computer-assisted diagnosis system for breast MRI. Journal of Magnetic Resonance Imaging*, 2012. **35**(5): p. 1077-1088.
 18. R. Altman, E. Asch, and D. Bloch, *Development of criteria for the classification and reporting of osteoarthritis. Classification of osteoarthritis of the knee. Arthritis and Rheumatism*, 1986. **29**(8): p. 1039-1052.
 19. J.D. Sipe, *Acute-phase proteins in osteoarthritis. Seminars in Arthritis and Rheumatism*, 1995. **25**(2): p. 75-86.
 20. K.A. Abbud Lozoya and J.B. Kouri Flores, *A novel rat osteoarthrosis model to assess apoptosis and matrix degradation. Pathology Research and Practice*, 2000. **196**(11): p. 729-745.
 21. M.B. Goldring, *The role of the chondrocyte in osteoarthritis. Arthritis and Rheumatism*, 2000. **43**(9): p. 1916-1926.
 22. S. Hashimoto, R.L. Ochs, F. Rosen, J. Quach, G. McCabe, J. Solan, J.E. Seegmiller, R. Terkeltaub, and M. Lotz, *Chondrocyte-derived apoptotic bodies*

- and calcification of articular cartilage*. Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(6): p. 3094-3099.
23. S. Rosengren, G.S. Firestein, and D.L. Boyle, *Measurement of Inflammatory Biomarkers in Synovial Tissue Extracts by Enzyme-Linked Immunosorbent Assay*. Clinical and Diagnostic Laboratory Immunology, 2003. **10**(6): p. 1002-1010.
 24. W.R. Hendee and C.J. Morgan, *Magnetic Resonance Imaging Part I-Physical Principles*. West J Med, 1984 **141**: p. 491-500.
 25. A. Wagner, M. Aurich, M. Stoessel, N. Sieber, W.D. Wetzel, J. Mollenhauer, K. Schmuck, C. Muehleman, M. Lohmann, B. Reime, J. Metge, P. Coan, A. Bravin, F. Arfelli, L. Rigon, and R.H. Menk. *Chance and limit of imaging of articular cartilage in vitro in healthy and arthritic joints - DEI (Diffraction Enhanced Imaging) in comparison with MRI, CT and ultrasound*. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2005. San Diego, CA.
 26. M. Cova and R. Toffanin, *MR microscopy of hyaline cartilage: Current status*. European Radiology, 2002. **12**(4): p. 814-823.
 27. F. Eckstein, D. Burstein, and T.M. Link, *Quantitative MRI of cartilage and bone: Degenerative changes in osteoarthritis*. NMR in Biomedicine, 2006. **19**(7): p. 822-854.
 28. G.E. Gold, B.A. Hargreaves, S.B. Reeder, S.S. Vasanawala, and C.F. Beaulieu, *Controversies in protocol selection in the imaging of articular cartilage*. Seminars in Musculoskeletal Radiology, 2005. **9**(2): p. 161-172.
 29. S.P.M. Dries, V. Pekar, D. Bystrov, H.S. Heese, T. Blaffert, C. Bos, and A.M.C. Van Muiswinkel. *Automatic knee cartilage delineation using inheritable*

- segmentation*. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2008. San Diego, CA.
30. J. Folkesson, E. Dam, O.F. Olsen, P. Pettersen, and C. Christiansen, *Automatic segmentation of the articular cartilage in knee MRI using a hierarchical multi-class classification scheme*. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, 2005. **8**(Pt 1): p. 327-334.
 31. J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin, *Automatic segmentation of articular cartilage in magnetic resonance images of the knee*. Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention, 2007. **10**(Pt 2): p. 186-194.
 32. A.A. Qazi, J. Folkesson, P.C. Pettersen, M.A. Karsdal, C. Christiansen, and E.B. Dam, *Separation of healthy and early osteoarthritis by automatic quantification of cartilage homogeneity*. Osteoarthritis and Cartilage, 2007. **15**(10): p. 1199-1206.
 33. A.A. Qazi, E.B. Dam, M. Loog, M. Nielsen, F. Lauze, and C. Christiansen. *A variational method for automatic localization of the most pathological ROI in the knee cartilage*. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2008. San Diego, CA.
 34. J.G. Tamez-Peña, M. Barbu-McInnis, and S. Totterman. *Structural quantification of cartilage changes using statistical parametric mapping*. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2007. San Diego, CA.

35. F. Eckstein, H.C. Charles, R.J. Buck, V.B. Kraus, A.E. Remmers, M. Hudelmaier, W. Wirth, and J.L. Evelhoch, *Accuracy and precision of quantitative assessment of cartilage morphology by magnetic resonance imaging at 3.0T*. Arthritis and Rheumatism, 2005. **52**(10): p. 3132-3136.
36. C. Glaser, *New techniques for cartilage imaging: T2 relaxation time and diffusion-weighted MR imaging*. Radiologic Clinics of North America, 2005. **43**(4): p. 641-653.
37. T.C. Dunn, Y. Lu, H. Jin, M.D. Ries, and S. Majumdar, *T2 relaxation time of cartilage at MR imaging: Comparison with severity of knee osteoarthritis*. Radiology, 2004. **232**(2): p. 592-598.
38. C. Liess, S. Lüsse, N. Karger, M. Heller, and C.G. Glüer, *Detection of changes in cartilage water content using MRI T2-mapping in vivo*. Osteoarthritis and Cartilage, 2002. **10**(12): p. 907-913.
39. X. Li, E.T. Han, D. Newitt, and S. Majumdar. *T1rho relaxation quantification using spiral imaging: A preliminary study*. in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. 2004. San Francisco, CA.
40. C. Taylor, J. Carballido-Gamio, S. Majumdar, and X. Li, *Comparison of quantitative imaging of cartilage for osteoarthritis: T2, T1ρ, dGEMRIC and contrast-enhanced computed tomography*. Magnetic Resonance Imaging, 2009. **27**(6): p. 779-784.
41. A.J. Wheaton, F.L. Casey, A.J. Gougoutas, G.R. Dodge, A. Borthakur, J.H. Lonner, H.R. Schumacher, and R. Reddy, *Correlation of T1ρ with fixed charge*

- density in cartilage*. Journal of Magnetic Resonance Imaging, 2004. **20**(3): p. 519-525.
42. G. Blumenkrantz and S. Majumdar, *Quantitative magnetic resonance imaging of articular cartilage in osteoarthritis*. European Cells and Materials, 2007. **13**: p. 75-86.
 43. D. Burstein, J. Velyvis, K.T. Scott, K.W. Stock, Y.J. Kim, D. Jaramillo, R.D. Boutin, and M.L. Gray, *Protocol issues for delayed Gd(DTPA)₂--enhanced MRI (dGEMRIC) for clinical evaluation of articular cartilage*. Magnetic Resonance in Medicine, 2001. **45**(1): p. 36-41.
 44. C.J. Tiderius, J. Tjörnstrand, P. Åkeson, K. Södersten, L. Dahlberg, and P. Leander, *Delayed gadolinium-enhanced MRI of cartilage (dGEMRIC): Intra-and interobserver variability in standardized drawing of regions of interest*. Acta Radiologica, 2004. **45**(6): p. 628-634.
 45. M. Hori, *Quantification of blood Gd-DTPA levels : implications on dosing dGEMRIC*, in *Massachusetts Institute of Technology. Dept. of Electrical Engineering and Computer Science*. 2005, Massachusetts Institute of Technology.
 46. S.E. Gwilym, J.R. Keltner, C.E. Warnaby, A.J. Carr, B. Chizh, I. Chessell, and I. Tracey, *Psychophysical and functional imaging evidence supporting the presence of central sensitization in a cohort of osteoarthritis patients*. Arthritis Care and Research, 2009. **61**(9): p. 1226-1234.
 47. H.M. Bassiouni, *Bone marrow lesions in the knee: The clinical conundrum*. International Journal of Rheumatic Diseases, 2010. **13**(3): p. 196-202.

48. M.E. Mayerhoefer, M. Breitsenseher, S. Hofmann, N. Aigner, R. Meizer, H. Siedentop, and J. Kramer, *Computer-assisted quantitative analysis of bone marrow edema of the knee: Initial experience with a new method*. American Journal of Roentgenology, 2004. **182**(6): p. 1399-1403.
49. C.G. Peterfy, A. Guermazi, S. Zaim, P.F.J. Tirman, Y. Miaux, D. White, M. Kothari, Y. Lu, K. Fye, S. Zhao, and H.K. Genant, *Whole-organ magnetic resonance imaging score (WORMS) of the knee in osteoarthritis*. Osteoarthritis and Cartilage, 2004. **12**(3): p. 177-190.
50. S. Hofmann, J. Kramer, M. Breitsenseher, M. Pietsch, and N. Aigner, *Bone marrow edema in the knee. Differential diagnosis and therapeutic possibilities*. Knochenmarködem im kniegelenk. Differenzialdiagnostik und therapeutische möglichkeiten, 2006. **35**(4): p. 463-477.
51. D.T. Felson, J. Niu, A. Guermazi, F. Roemer, P. Aliabadi, M. Clancy, J. Torner, C.E. Lewis, and M.C. Nevitt, *Correlation of the development of knee pain with enlarging bone marrow lesions on magnetic resonance imaging*. Arthritis and Rheumatism, 2007. **56**(9): p. 2986-2992.
52. D.J. Hunter, Y. Zhang, J. Niu, J. Goggins, S. Amin, M.P. LaValley, A. Guermazi, H. Genant, D. Gale, and D.T. Felson, *Increase in bone marrow lesions associated with cartilage loss: A longitudinal magnetic resonance imaging study of knee osteoarthritis*. Arthritis & Rheumatism, 2006. **54**(5): p. 1529-1535.
53. F.W. Roemer, A. Guermazi, M.K. Javaid, J.A. Lynch, J. Niu, Y. Zhang, D.T. Felson, C.E. Lewis, J. Torner, and M.C. Nevitt, *Change in MRI-detected subchondral bone marrow lesions is associated with cartilage loss: The MOST*

- Study. A longitudinal multicentre study of knee osteoarthritis. Annals of the Rheumatic Diseases*, 2009. **68**(9): p. 1461-1465.
54. D.T. Felson, S. McLaughlin, J. Goggins, M.P. LaValley, M.E. Gale, S. Totterman, W. Li, C. Hill, and D. Gale, *Bone marrow edema and its relation to progression of knee osteoarthritis. Annals of internal medicine*, 2003. **139**(5 Pt 1): p. 330-336.
 55. J. Shiraishi, Q. Li, D. Appelbaum, and K. Doi, *Computer-aided diagnosis and artificial intelligence in clinical imaging. Seminars in Nuclear Medicine*, 2011. **41**(6): p. 449-462.
 56. T. Stammberger, F. Eckstein, M. Michaelis, K.H. Englmeier, and M. Reiser, *Interobserver reproducibility of quantitative cartilage measurements: Comparison of B-spline snakes and manual segmentation. Magnetic Resonance Imaging*, 1999. **17**(7): p. 1033-1042.
 57. D.L. Pham, C. Xu, and J.L. Prince, *Current methods in medical image segmentation Annual Review of Biomedical Engineering* 2000. **2**: p. 315-337.
 58. Y. Sun, E.C. Teo, and Q.H. Zhang. *Discussions of knee joint segmentation. in ICBPE 2006 - Proceedings of the 2006 International Conference on Biomedical and Pharmaceutical Engineering*. 2006. Singapore.
 59. J.P.W. Pluim, J.B.A. Maintz, and M.A. Viergever, *Mutual-information-based registration of medical images: A survey. IEEE Transactions on Medical Imaging*, 2003. **22**(8): p. 986-1004.
 60. B. Zitová and J. Flusser, *Image registration methods: A survey. Image and Vision Computing*, 2003. **21**(11): p. 977-1000.

61. M.S. Aslan, E. Mostafa, H. Abdelmunim, A. Shalaby, A.A. Farag, and B. Arnold. *A novel probabilistic simultaneous segmentation and registration using level set*. 2011. Brussels.
62. C. Lu, J. Zhu, and J.S. Duncan. *An algorithm for simultaneous image segmentation and nonrigid registration, with clinical application in image guided radiotherapy*. 2011. Brussels.
63. S. Mizuta, S.I. Urayama, H. Watabe, N. Sugimoto, and C. Uyama. *Simultaneous process of automated 3D registration and segmentation on medical images*. 1997. Newport Beach, CA.
64. S. Ghosh, O. Beuf, M. Ries, N.E. Lane, L.S. Steinbach, T.M. Link, and S. Majumdar. *Watershed segmentation of high resolution magnetic resonance images of articular cartilage of the knee*. in *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*. 2000. Chicago, IL.
65. E.B. Dam, J. Folkesson, P.C. Pettersen, and C. Christiansen, *Semi-automatic knee cartilage segmentation*. *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*, 2006. **6144 II**.
66. J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin, *Automatic segmentation of articular cartilage in magnetic resonance images of the knee*. *MICCAI*, 2007. **10**(Pt 2): p. 186-194.
67. Part, J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin, *Automatic segmentation of articular cartilage in magnetic resonance images of the knee*, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2007: Brisbane. p. 186-194.

68. H. Shim, S. Lee, B. Kim, C. Tao, S. Chang, I.D. Yun, S.U. Lee, K. Kwoh, and K. Bae. *3-D segmentation of articular cartilages by graph cuts using knee MR images from the osteoarthritis initiative*. in *Progress in Biomedical Optics and Imaging - Proceedings of SPIE*. 2008. San Diego, CA.
69. P. Dodin, J.P. Pelletier, J. Martel-Pelletier, and F. Abram, *Automatic human knee cartilage segmentation from 3-D magnetic resonance images*. IEEE Transactions on Biomedical Engineering, 2010. **57**(11): p. 2699-2711.
70. A.B. Wolff, L.L. Pesce, J.S. Wu, L.R. Smart, M.J. Medvecky, and A.H. Haims, *Comparison of spin echo T1-weighted sequences versus fast spin-echo proton density-weighted sequences for evaluation of meniscal tears at 1.5 T*. Skeletal Radiology, 2009. **38**(1): p. 21-29.
71. J. Fripp, S. Crozier, S.K. Warfield, and S. Ourselin, *Automatic segmentation of the bone and extraction of the bone-cartilage interface from magnetic resonance images of the knee*. Physics in Medicine and Biology, 2007. **52**(6): p. 1617-1631.
72. H. Seim, D. Kainmueller, H. Heller, H. Lamecker, S. Zachow, and H.C. Hege, *Automatic Segmentation of the Pelvic Bones from CT Data Based on a Statistical Shape Model*. Eurographics Workshop on Visual Computing for Biomedicine, 2008: p. 93-100.
73. J. Fripp, P. Bourgeat, A.J.U. Mewes, S.K. Warfield, S. Crozier, and S. Ourselin, *3D statistical shape models to embed spatial relationship information*, in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*. 2005: Beijing. p. 51-60.

74. H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, *Model-based Auto-Segmentation of Knee Bones and Cartilage in MRI Data*. MICCAI 2010 Workshop, 2010: p. 215 - 223
75. M. Kass, A. Witkin, and D. Terzopoulos, *Snakes: Active contour models*. International Journal of Computer Vision, 1988. **1**(4): p. 321-331.
76. M.S. Horritt, D.C. Mason, and A.J. Luckman, *Flood boundary delineation from synthetic aperture radar imagery using a statistical active contour model*. International Journal of Remote Sensing, 2001. **22**(13): p. 2489-2507.
77. H. Jiang and Q. Cheng. *Automatic 3D segmentation of CT images based on active contour models*. in *11th IEEE International Conference on Computer-Aided Design and Computer Graphics*. 2009. Huangshan.
78. V. Jurcak, J. Fripp, C. Engstrom, D. Walker, O. Salvado, S. Ourselin, and S. Crozier, *Automated segmentation of the quadratus lumborum muscle from magnetic resonance images using a hybrid atlas based - geodesic active contour scheme*. Conference proceedings : Annual International Conference of the IEEE EMBS., 2008. **2008**: p. 867-870.
79. B. Liu, H.D. Cheng, J. Huang, J. Tian, J. Liu, and X. Tang, *Automated Segmentation of Ultrasonic Breast Lesions Using Statistical Texture Classification and Active Contour Based on Probability Distance*. Ultrasound in Medicine and Biology, 2009. **35**(8): p. 1309-1324.
80. T.F. Chan and L.A. Vese, *Active contours without edges*. IEEE Transactions on Image Processing, 2001. **10**(2): p. 266-277.

81. L.D. Cohen, *On active contour models and balloons*. CVGIP: Image Understanding, 1991. **53**(2): p. 211-218.
82. C. Xu and J.L. Prince, *Snakes, shapes, and gradient vector flow*. IEEE Transactions on Image Processing, 1998. **7**(3): p. 359-369.
83. N. Ray, S.T. Acton, T. Altes, E.E. De Lange, and J.R. Brookeman, *Merging parametric active contours within homogeneous image regions for MRI-based lung segmentation*. IEEE Transactions on Medical Imaging, 2003. **22**(2): p. 189-199.
84. T. McInerney and D. Terzopoulos, *T-snakes: Topology adaptive snakes*. Medical Image Analysis, 2000. **4**(2): p. 73-91.
85. C. Li, C. Xu, C. Gui, and M.D. Fox. *Level set evolution without re-initialization: A new variational formulation*. 2005. San Diego, CA.
86. H. Zhou, C. Yuan, and Y. Liu. *Fast level set-based segmentation without edge and re-initialization*. 2010. Innsbruck.
87. L.A. Vese and T.F. Chan, *A multiphase level set framework for image segmentation using the Mumford and Shah model*. International Journal of Computer Vision, 2002. **50**(3): p. 271-293.
88. J. Jiang, Y. Guo, S. Zhan, and H. Li. *Segmentation of knee joints based on improved multiphase Chan-Vese model*. 2008. Shanghai.
89. J. Carballido-Gamio, J.S. Bauer, R. Stahl, K.Y. Lee, S. Krause, T.M. Link, and S. Majumdar, *Inter-subject comparison of MRI knee cartilage thickness*. Medical Image Analysis, 2008. **12**(2): p. 120-135.

90. T.G. Williams, A.P. Holmes, M. Bowes, G. Vincent, C.E. Hutchinson, J.C. Waterton, R.A. Maciewicz, and C. Taylor, *Measurement and visualisation of focal cartilage thickness change by MRI in a study of knee osteoarthritis using a novel image analysis tool*. British Journal of Radiology, 2010. **83**(995): p. 940-948.
91. A.A. Qazi, E.B. Dam, M. Nielsen, M.A. Karsdal, P.C. Pettersen, and C. Christiansen, *Osteoarthritic Cartilage Is More Homogeneous Than Healthy Cartilage. Identification of a Superior Region of Interest Colocalized With a Major Risk Factor for Osteoarthritis*. Academic Radiology, 2007. **14**(10): p. 1209-1220.
92. A.J. Dijkstra, P. Anbeek, K.G. Auw Yang, K.L. Vincken, M.A. Viergever, R.M. Castelein, and D.B.F. Saris, *Validation of a Novel Semiautomated Segmentation Method for MRI Detection of Cartilage-Related Bone Marrow Lesions*. Cartilage, 2010. **1**(4): p. 328-334.
93. R. Burgkart, C. Glaser, A. Hyhlik-Dürr, K.H. Englmeier, M. Reiser, and F. Eckstein, *Magnetic Resonance Imaging-Based Assessment of Cartilage Loss in Severe Osteoarthritis: Accuracy, Precision, and Diagnostic Value*. Arthritis and Rheumatism, 2001. **44**(9): p. 2072-2077.
94. T. Stammberger, F. Eckstein, K.H. Englmeier, and M. Reiser, *Determination of 3D cartilage thickness data from MR imaging: Computational method and reproducibility in the living*. Magnetic Resonance in Medicine, 1999. **41**(3): p. 529-536.

95. A.A. Kshirsagar, P.J. Watson, J.A. Tyler, and L.D. Hall, *Measurement of localized cartilage volume and thickness of human knee joints by computer analysis of three-dimensional magnetic resonance images*. Investigative Radiology, 1998. **33**(5): p. 289-299.
96. O.J. Muensterer, F. Eckstein, D. Hahn, and R. Putz, *Computer-aided three dimensional assessment of knee-joint cartilage with magnetic resonance imaging*. Clinical Biomechanics, 1996. **11**(5): p. 260-266.
97. M. Haubner, F. Eckstein, M. Schnier, A. Losch, H. Sittek, C. Becker, H. Kolem, M. Reiser, and K.H. Englmeier, *A non-invasive technique for 3-dimensional assessment of articular cartilage thickness based on MRI .2. Validation using CT arthrography*. Magnetic Resonance Imaging, 1997. **15**(7): p. 805-813.
98. S. Akhtar, C.L. Poh, and R.I. Kitney, *An MRI derived articular cartilage visualization framework*. Osteoarthritis and Cartilage, 2007. **15**(9): p. 1070-1085.
99. P.M.M. Cashman, R.I. Kitney, M.A. Gariba, and M.E. Carter, *Automated techniques for visualization and mapping of articular cartilage in MR images of the osteoarthritic knee: A base technique for the assessment of microdamage and submicro damage*. IEEE Transactions on Nanobioscience, 2002. **1**(1): p. 42-51.
100. M.E. Bowers, N. Trinh, G.A. Tung, J.J. Crisco, B.B. Kimia, and B.C. Fleming, *Quantitative MR imaging using "LiveWire" to measure tibiofemoral articular cartilage thickness*. Osteoarthritis and Cartilage, 2008. **16**(10): p. 1167-1173.
101. J. Carballido-Gamio, T.M. Link, and S. Majumdar, *New techniques for cartilage magnetic resonance imaging relaxation time analysis: Texture analysis of*

- flattened cartilage and localized intra- and inter-subject comparisons. Magnetic Resonance in Medicine*, 2008. **59**(6): p. 1472-1477.
102. T.M. Link, S. Majumdar, P. Augat, J.C. Lin, D. Newitt, Y. Lu, N.E. Lane, and H.K. Genant, *In vivo high resolution MRI of the calcaneus: Differences in trabecular structure in osteoporosis patients. Journal of Bone and Mineral Research*, 1998. **13**(7): p. 1175-1182.
 103. D. Mueller, T.M. Link, R. Monetti, J. Bauer, H. Boehm, V. Seifert-Klauss, E.J. Rummeny, G.E. Morfill, and C. Raeth, *The 3D-based scaling index algorithm: A new structure measure to analyze trabecular bone architecture in high-resolution MR images in vivo. Osteoporosis International*, 2006. **17**(10): p. 1483-1493.
 104. H.F. Boehm, T.M. Link, R.A. Monetti, D. Mueller, E.J. Rummeny, D. Newitt, S. Majumdar, and C.W. Raeth. *Application of the minkowski functionals in 3D to high resolution MR images of trabecular bone: Prediction of the biomechanical strength by non-linear topological measures. in Medical Imaging 2004: Imaging Processing*. 2004. San Diego, CA.
 105. H.F. Boehm, C. Raeth, R.A. Monetti, D. Mueller, D. Newitt, S. Majumdar, E. Rummeny, G. Morfill, and T.M. Link, *Local 3D scaling properties for the analysis of trabecular bone extracted from high-resolution magnetic resonance imaging of human trabecular bone: Comparison with bone mineral density in the prediction of biomechanical strength in vitro. Investigative Radiology*, 2003. **38**(5): p. 269-280.
 106. H.F. Boehm, C.W. Raeth, R. Monetti, D. Mueller, D. Newitt, S. Majumdar, E.J. Rummeny, and T.M. Link. *Application of the standard hough-transform to high*

- resolution MRI of human trabecular bone to predict mechanical strength*. 2003. San Diego, CA.
107. T.M. Link, V. Vieth, R. Langenberg, N. Meier, A. Lotter, D. Newitt, and S. Majumdar, *Structure analysis of high resolution magnetic resonance imaging of the proximal femur: In vitro correlation with biomechanical strength and BMD*. *Calcified Tissue International*, 2003. **72**(2): p. 156-165.
 108. R.A. Monetti, H. Böhm, D. Müller, E. Rummeny, T. Link, and C. Rth. *Structural analysis of human proximal femur for the prediction of biomechanical strength 'in vitro'. The locally adapted scaling vector method*. 2005. San Diego, CA.
 109. C.W. Raeth, D. Mueller, E.J. Rummeny, T.M. Link, M. Vogel, H. Koenig, H. Boehm, and R. Monetti. *Optimizing texture measures quantifying bone structures as well as MR-sequences at 3 tesla: An integrative statistical approach*. 2006. San Diego, CA.
 110. N. Boutry, B. Cortet, P. Dubois, X. Marchandise, and A. Cotten, *Trabecular bone structure of the calcaneus: Preliminary in vivo MR imaging assessment in men with osteoporosis*. *Radiology*, 2003. **227**(3): p. 708-717.
 111. G. Castellano, L. Bonilha, L.M. Li, and F. Cendes, *Texture analysis of medical images*. *Clinical Radiology*, 2004. **59**(12): p. 1061-1069.
 112. M.E. Mayerhoefer, M.J. Breitenseher, J. Kramer, N. Aigner, S. Hofmann, and A. Materka, *Texture analysis for tissue discrimination on T1-weighted MR images of the knee joint in a multicenter study: Transferability of texture features and comparison of feature selection methods and classifiers*. *Journal of Magnetic Resonance Imaging*, 2005. **22**(5): p. 674-680.

113. M. Khider, A. Taleb-Ahmed, P. Dubois, and B. Haddad. *Classification of trabecular bone texture from MRI and CT scan images by multi resolution analysis*. in *29th Annual International Conference of IEEE-EMBS, Engineering in Medicine and Biology Society, EMBC'07*. 2007. Lyon.
114. K. Brandt and C. Slemenda, *Osteoarthritis: epidemiology, pathology and pathogenesis*. 10th ed. Schumacher HR, Ed. Primer on the Rheumatic Diseases. 1993: Arthritis Foundation. 184-188.
115. M.P. Recht, D.W. Goodwin, C.S. Winalski, and L.M. White, *MRI of articular cartilage: Revisiting current status and future directions*. American Journal of Roentgenology, 2005. **185**(4): p. 899-914.
116. C.G. Peterfy, *Role of MR imaging in clinical research studies*. Seminars in Musculoskeletal Radiology, 2001. **5**(4): p. 365-378.
117. F. Eckstein and C. Glaser, *Measuring cartilage morphology with quantitative magnetic resonance imaging*. Seminars in Musculoskeletal Radiology, 2004. **8**(4): p. 329-353.
118. C. Buckland-Wright, *Current status of imaging procedures in the diagnosis, prognosis and monitoring of osteoarthritis*. Bailliere's Clinical Rheumatology, 1997. **11**(4): p. 727-748.
119. L.P. Clarke, R.P. Velthuisen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, and M.L. Silbiger, *MRI segmentation: Methods and applications*. Magnetic Resonance Imaging, 1995. **13**(3): p. 343-368.
120. J.P. Raynauld, *Quantitative magnetic resonance imaging of articular cartilage in knee osteoarthritis*. Current Opinion in Rheumatology, 2003. **15**(5): p. 647-650.

121. H. Imhof, I.M. Nöbauer-Huhmann, C. Krestan, A. Gahleitner, I. Sulzbacher, S. Marlovits, and S. Trattnig, *MRI of the cartilage*. European Radiology, 2002. **12**(11): p. 2781-2793.
122. C.A. Helms, *Fundamentals of Skeletal Radiology* 2005, Elsevier Health Sciences. p. 235.
123. J.P. Raynauld, J. Martel-Pelletier, M.J. Berthiaume, G. Beaudoin, D. Choquette, B. Haraoui, H. Tannenbaum, J.M. Meyer, J.F. Beary, G.A. Cline, and J.P. Pelletier, *Long term evaluation of disease progression through the quantitative magnetic resonance imaging of symptomatic knee osteoarthritis patients: Correlation with clinical symptoms and radiographic changes*. Arthritis Research and Therapy, 2006. **8**(1).
124. J.P. Raynauld, J. Martel-Pelletier, M.J. Berthiaume, F. Labonte, G. Beaudoin, J.A. de Guise, D.A. Bloch, D. Choquette, B. Haraoui, R.D. Altman, M.C. Hochberg, J.M. Meyer, G.A. Cline, and J.P. Pelletier, *Quantitative magnetic resonance imaging evaluation of knee osteoarthritis progression over two years and correlation with clinical symptoms and radiologic changes*. Arthritis and Rheumatism, 2004. **50**(2): p. 476-487.
125. K.Y. Lee, T.C. Dunn, L.S. Steinbach, E. Ozhinsky, M.D. Ries, and S. Majumdar, *Computer-aided quantification of focal cartilage lesions of osteoarthritic knee using MRI*. Magnetic Resonance Imaging, 2004. **22**(8): p. 1105-1115.
126. C.T. Lindsey, A. Narasimhan, J.M. Adolfo, H. Jin, L.S. Steinbach, T. Link, M. Ries, and S. Majumdar, *Magnetic resonance evaluation of the interrelationship*

- between articular cartilage and trabecular bone of the osteoarthritic knee.*
- Osteoarthritis and Cartilage, 2004. **12**(2): p. 86-96.
127. M.P. Recht, D.W. Piraino, G.A. Paletta, J.P. Schils, and G.H. Belhobek, *Accuracy of fat-suppressed three-dimensional spoiled gradient-echo FLASH MR imaging in the detection of patellofemoral articular cartilage abnormalities.* Radiology, 1996. **198**(1): p. 209-212.
 128. C. Kauffmann, P. Gravel, B. Godbout, A. Gravel, G. Beaudoin, J.P. Raynauld, J. Martel-Pelletier, J.P. Pelletier, and J.A. de Guise, *Computer-aided method for quantification of cartilage thickness and volume changes using MRI: validation study using a synthetic model.* Biomedical Engineering, IEEE Transactions on, 2003. **50**(8): p. 978-988.
 129. F. Eckstein, F. Cicuttini, J.P. Raynauld, J.C. Waterton, and C. Peterfy, *Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): morphological assessment.* Osteoarthritis and Cartilage, 2006. **14**(SUPPL. 1): p. 46-75.
 130. A.H. Sonin, R.A. Pensy, M.E. Mulligan, and S. Hatem, *Grading articular cartilage of the knee using fast spin-echo proton density-weighted MR imaging without fat suppression.* American Journal of Roentgenology, 2002. **179**(5): p. 1159-1166.
 131. A. Kolk, C. Pautke, E. Wiener, V. Schott, K.-D. Wolff, H.-H. Horch, and E.J. Rummeny, *Isotropic proton-density-weighted high-resolution MRI for volume measurement of reconstructed orbital fractures -- a comparison with multislice CT.* Magnetic Resonance Imaging, 2008. **26**(8): p. 1167-1174.

132. S.-Y. Lee, W.-H. Jee, S. Kim, and J.-M. Kim, *Proton density-weighted MR imaging of the knee: fat suppression versus without fat suppression*. Skeletal Radiology, 2010: p. 1-7.
133. J. Hohe, S. Faber, R. Muehlbauer, M. Reiser, K.H. Englmeier, and F. Eckstein, *Three-dimensional analysis and visualization of regional MR signal intensity distribution of articular cartilage*. Medical Engineering and Physics, 2002. **24**(3): p. 219-227.
134. W.S. Rasband. *ImageJ*. <http://rsb.info.nih.gov/ij/> 1997-2008; 1.43u:[Available from: <http://rsb.info.nih.gov/ij/>].
135. Y. Sivan. *gaussian curve fit*. 2006 19 Jul 2006; Available from: <http://www.mathworks.com/matlabcentral/fileexchange/11733-gaussian-curve-fit>.
136. SSfM. *LSGE: The Least Squares Geometric Elements library consists of MatLab functions to find ...lscylinder*. 2004 [cited 2010; Available from: http://scicomp.npl.co.uk/eurometros/gen_report.php?category=distributions&pkey=14&subform=yes].
137. J.H. Ho, W.Z. Lung, C.L. Sean, C.L. Poh, K. Sheah, D.T.T. Lie, and K.S.A. Yew. *Anterior cruciate ligament segmentation using morphological operations with active contour*. 2010. Chengdu.
138. A. Losch, F. Eckstein, M. Haubner, and K.H. Englmeier, *A non-invasive technique for 3-dimensional assessment of articular cartilage thickness based on MRI .I. Development of a computational method*. Magnetic Resonance Imaging, 1997. **15**(7): p. 795-804.

139. Z.A. Cohen, V.C. Mow, J.H. Henry, W.N. Levine, and G.A. Ateshian, *Templates of the cartilage layers of the patellofemoral joint and their use in the assessment of osteoarthritic cartilage damage*. Osteoarthritis and Cartilage, 2003. **11**(8): p. 569-579.
140. H.G. Potter, J.M. Linklater, A.A. Allen, J.A. Hannafin, and S.B. Haas, *Magnetic resonance imaging of articular cartilage in the knee. An evaluation with use of fast-spin-echo imaging*. Journal of Bone and Joint Surgery - Series A, 1998. **80**(9): p. 1276-1284.
141. F.W. Roemer, R. Frobell, D.J. Hunter, M.D. Crema, W. Fischer, K. Bohndorf, and A. Guermazi, *MRI-detected subchondral bone marrow signal alterations of the knee joint: terminology, imaging appearance, relevance and radiological differential diagnosis*. Osteoarthritis and Cartilage, 2009. **17**(9): p. 1115-1131.
142. C.G. Peterfy, *MRI of the wrist in early rheumatoid arthritis*. Annals of the Rheumatic Diseases, 2004. **63**(5): p. 473-477.
143. A. Bernasconi, S.B. Antel, D.L. Collins, N. Bernasconi, A. Olivier, F. Dubeau, G.B. Pike, F. Andermann, and D.L. Arnold, *Texture analysis and morphological processing of magnetic resonance imaging assist detection of focal cortical dysplasia in extra-temporal partial epilepsy*. Annals of Neurology, 2001. **49**(6): p. 770-775.
144. J.M. Mathias, P.S. Tofts, and N.A. Losseff, *Texture analysis of spinal cord pathology in multiple sclerosis*. Magnetic Resonance in Medicine, 1999. **42**(5): p. 929-935.

145. S. Sinha, F.A. Lucas-Quesada, N.D. DeBruhl, J. Sayre, D. Farria, D.P. Gorczyca, and L.W. Bassett, *Multifeature analysis of Gd-enhanced MR images of breast lesions*. Journal of Magnetic Resonance Imaging, 1997. **7**(6): p. 1016-1026.
146. Q. Ji, J. Engel, and E. Craine, *Texture analysis for classification of cervix lesions*. IEEE Transactions on Medical Imaging, 2000. **19**(11): p. 1144-1149.
147. N.J. Tustison, B.B. Avants, P.A. Cook, Y. Zheng, A. Egan, P.A. Yushkevich, and J.C. Gee, *N4ITK: Improved N3 bias correction*. IEEE Transactions on Medical Imaging, 2010. **29**(6): p. 1310-1320.
148. P.M. Szczypiński, M. Strzelecki, A. Materka, and A. Klepaczko, *MaZda-A software package for image texture analysis*. Computer Methods and Programs in Biomedicine, 2009. **94**(1): p. 66-76.
149. A. Materka. *MaZda User's Manual*.
http://www.eletel.p.lodz.pl/programy/mazda/download/mazda_manual.pdf 1998-2005 [1 December 2011]; Available from:
http://www.eletel.p.lodz.pl/programy/mazda/download/mazda_manual.pdf.
150. R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax, and S. Verzakov, *A Matlab Toolbox for Pattern Recognition*. 2007. **Delft University of Technology**.
151. J. Schürmann, *Pattern Classification: A Unified View of Statistical and Neural Approaches*. 1996: John Wiley & Sons, Inc.
152. P.C. Mahalanobis, *On the generalised distance in statistics*. Proceedings of the National Institute of Sciences of India 1936. **2**(1): p. 49-55.

153. M.M. Galloway, *Texture Analysis Using Gray Level Run Lengths*. Computer Graphics and Image Processing, 1975. **4**: p. 172-179.
154. A.W. Whitney, *A Direct Method of Nonparametric Measurement Selection*. Computers, IEEE Transactions on, 1971. **C-20**(9): p. 1100-1103.
155. T. Fawcett, *An introduction to ROC analysis*. Pattern Recognition Letters, 2006. **27**(8): p. 861-874.
156. S.T. Acton and N. Ray, *Algorithm KWT*, in *Biomedical Image Analysis: Segmentation*, A.C. Bovik, Editor. 2009, Morgan & Claypool Publishers. p. 11-12.
157. Y. Chen, W. Zhao, and Z. Wang. *Level set segmentation algorithm based on image entropy and simulated annealing*. in *ICBBE*. 2007. Wuhan.
158. S.H. Lee and J.K. Seo, *Level set-based bimodal segmentation with stationary global minimum*. IEEE Transactions on Image Processing, 2006. **15**(9): p. 2843-2852.
159. N. Ray, B. Chanda, and J. Das, *A fast and flexible multiresolution snake with a definite termination criterion*. Pattern Recognition, 2001. **34**(7): p. 1483-1490.
160. Y. Shuai, H. Sun, and G. Xu, *SAR image segmentation based on level set with stationary global minimum*. IEEE Geoscience and Remote Sensing Letters, 2008. **5**(4): p. 644-648.
161. X.F. Wang, D.S. Huang, and H. Xu, *An efficient local Chan-Vese model for image segmentation*. Pattern Recognition, 2010. **43**(3): p. 603-618.
162. J. Liu, W. Gao, S. Huang, and W.L. Nowinski, *A model-based, semi-global segmentation approach for automatic 3-D point landmark localization in neuroimages*. IEEE Transactions on Medical Imaging, 2008. **27**(8): p. 1034-1044.

163. S. Lobregt and M.A. Viergever, *Discrete dynamic contour model*. IEEE Transactions on Medical Imaging, 1995. **14**(1): p. 12-24.
164. X. Dong and G. Zheng, *Automatic extraction of proximal femur contours from calibrated x-ray images using 3D statistical models: An in vitro study*. International Journal of Medical Robotics and Computer Assisted Surgery, 2009. **5**(2): p. 213-222.
165. M.A. Baldwin, J.E. Langenderfer, P.J. Rullkoetter, and P.J. Laz, *Development of subject-specific and statistical shape models of the knee using an efficient segmentation and mesh-morphing approach*. Computer Methods and Programs in Biomedicine, 2010. **97**(3): p. 232-240.
166. U. Von Jan, D. Sandkühler, L. Kirsch, U. Rühmann, and H.M. Overhoff, *Surgical planning for the humeral part of a shoulder endoprosthesis based on landmarks determined from 3D ultrasound volumes*. International Journal of Computer Assisted Radiology and Surgery, 2006. **1**(SUPPL. 7): p. 229-231.
167. N. Otsu, *Threshold selection method from gray-level histograms*. IEEE Trans Syst Man Cybern, 1979. **SMC-9**(1): p. 62-66.
168. L. Lam, S.-W. Lee, and C.Y. Suen, *Thinning methodologies--A comprehensive survey*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992. **14**(9): p. page 879, bottom of first column through top of second column.
169. S. Jehan-Besson, M. Gastaud, M. Barlaud, and G. Aubert. *Region-based active contours using geometrical and statistical features for image segmentation*. in *IEEE International Conference on Image Processing*. 2003. Barcelona.

170. D. Mueller, R.A. Monetti, H.F. Boehm, J. Bauer, E.J. Rummeny, T.M. Link, and C.W. Raeth, *The 3D-based scaling index algorithm to optimize structure analysis of trabecular bone in postmenopausal women with and without osteoporotic spine fractures*. in *Proceedings of SPIE*. 2004. San Diego, CA.
171. M. Wolski, P. Podsiadlo, and G.W. Stachowiak, *Directional fractal signature analysis of trabecular bone: evaluation of different methods to detect early osteoarthritis in knee radiographs*. *Proceedings of the Institution of Mechanical Engineers. Part H, Journal of engineering in medicine*, 2009. **223**(2): p. 211-236.
172. X. Han, M.S. Hoogeman, P.C. Levendag, L.S. Hibbard, D.N. Teguh, P. Voet, A.C. Cowen, and T.K. Wolf, *Atlas-based auto-segmentation of head and neck CT images*, in *Lecture Notes in Computer Science* 2008: New York, NY. p. 434-441.
173. T. Rohlfing, R. Brandt, R. Menzel, D. Russakoff, and C. Maurer Jr, *Quo vadis, atlas-based segmentation?* *The Handbook of Medical Image Analysis*, 2005: p. 435-486.
174. R. Ronfard, *Region-based strategies for active contour models*. *International Journal of Computer Vision*, 1994. **13**(2): p. 229-251.
175. C. Tauber, H. Batatia, and A. Ayache, *Quasi-automatic initialization for parametric active contours*. *Pattern Recognition Letters*, 2010. **31**(1): p. 83-90.
176. M.B. Cuadra, C. Polio, A. Bardera, O. Cuisenaire, J.G. Villemure, and J.P. Thiran, *Atlas-based segmentation of pathological MR brain images using a model of lesion growth*. *IEEE Transactions on Medical Imaging*, 2004. **23**(10): p. 1301-1314.

177. Able Software Corp. *3D-Doctor 4.0.200800808*. 1994-2009 [cited 2009 23rd Oct 2009]; Available from: <http://www.ablesw.com/3d-doctor/publist.html>.
178. P. Heinze, D. Meister, R. Kober, J. Raczowsky, and H. Wörn, *Atlas-based segmentation of pathological knee joints*. Studies in health technology and informatics, 2002. **85**: p. 198-203.
179. N.M. Alpert, J.F. Bradshaw, D. Kennedy, and J.A. Correia, *The principal axes transformation-A method for image registration*. Journal of Nuclear Medicine, 1990. **31**(10): p. 1717-1722.
180. T. Chuah, J. Lim, and C. Poh. *Group Average Difference: A Termination Criterion for Active Contour*. Journal of Digital Imaging, **2012**: **25**(2) p. 279-93.