

# Automatic Segmentation of Articular Cartilage in Magnetic Resonance Images of the Knee

Jurgen Frupp<sup>1,2</sup>, Stuart Crozier<sup>2</sup>, Simon K. Warfield<sup>3</sup>, and Sébastien Ourselin<sup>1</sup>

<sup>1</sup> BioMedIA Lab, e-Health Research Centre  
CSIRO ICT Centre, Australia

{jurgen.frupp, sebastien.ourselin}@csiro.au

<sup>2</sup> School of ITEE, University of Queensland, Australia  
stuart@itee.uq.edu.au

<sup>3</sup> Computational Radiology Laboratory, Harvard Medical School,  
Children's Hospital Boston, USA  
warfield@crl.med.harvard.edu

**Abstract.** To perform cartilage quantitative analysis requires the accurate segmentation of each individual cartilage. In this paper we present a model based scheme that can automatically and accurately segment each individual cartilage in healthy knees from a clinical MR sequence (fat suppressed spoiled gradient recall). This scheme consists of three stages; the automatic segmentation of the bones, the extraction of the bone-cartilage interfaces (BCI) and segmentation of the cartilages. The bone segmentation is performed using three-dimensional active shape models. The BCI is extracted using image information and prior knowledge about the likelihood of each point belonging to the interface. A cartilage thickness model then provides constraints and regularizes the cartilage segmentation performed from the BCI. The accuracy and robustness of the approach was experimentally validated, with (patellar, tibial and femoral) cartilage segmentations having a median DSC of (0.870, 0.855, 0.870), performing significantly better than non-rigid registration (0.787, 0.814, 0.795). The total cartilage segmentation had an average DSC of (0.891), close to the (0.896) obtained using a semi-automatic watershed algorithm. The error in quantitative volume and thickness measures was (8.29, 4.94, 5.56)% and (0.19, 0.33, 0.10) *mm* respectively.

## 1 Introduction

MR imaging allows the non-invasive assessment of cartilage tissue, which is required for clinical studies, surgical treatments and drug trials. To obtain statistical significance, this assessment is usually performed on each cartilage separately (or in subregions) using morphological measures (volume, thickness, surface area or curvature) [1]. These measures require the cartilage to be segmented separately (or in subregions), a task that can significantly influence the error and reproducibility of the quantitative analysis, and has thus far proved difficult to automate, with current approaches manual [2] or semi-automatic (e.g. region growing [3], B-spline snakes [4] and live-wires [5]). These approaches take 30

minutes to several hours and require a skilled trained operator to obtain reproducible results.

$T_1$  weighted fat suppressed (FS) spoiled gradient recall images (SPGR) are the most commonly used MR sequence for cartilage assessment. They have poor cartilage-meniscus and cartilage-synovial fluid contrast and can exhibit significant imaging artifacts that obscure and artificially create defects [6]. These along with image resolution, magic angle and partial volume effects cause the problems that make accurate cartilage segmentation difficult. A few recently developed sequences reduce many of these problems [7]. In diseased knees, osteophytes, lesions and cracks are commonly observed and make segmentation more difficult.

The development of an automatic approach is desirable and has been pursued by several groups. With healthy knees, accurate segmentation results (Dice Similarity Coefficient (DSC)  $\approx 0.9$ ) were obtained by Grau using a modified semi-automatic watershed metric that utilizes prior information and requires around 10 minutes of user seeding [8]. Unfortunately, it segments all the individual cartilages as a single object, which is not sensitive enough for use in quantitative analysis. Using low field non-FS  $T_1$ -weighted MR images, a tissue classifier has been found to obtain reasonably accurate results on a large database of healthy and OA patients (DSC  $\approx 0.8$ ) [9]. In [10] an almost fully automatic approach has been used to segment ankle cartilages by first segmenting the bones, creating a surface mesh from which a local graph is built and two (bone and cartilage) surfaces are extracted simultaneously using two separate cost functions.

In this paper we present a model based scheme that uses localization obtained from a segmentation hierarchy as well as prior knowledge of the cartilage position and thickness variation to accurately segment each individual cartilage (patellar, tibial and femoral) in healthy knees. This scheme was validated and compared with FFD based NRR and the modified watershed approach of Grau [8]. The influence of repositioning and partial voluming effects was evaluated. Finally, we demonstrate the promise of this approach using volume and thickness quantitative measures.

## 2 Method

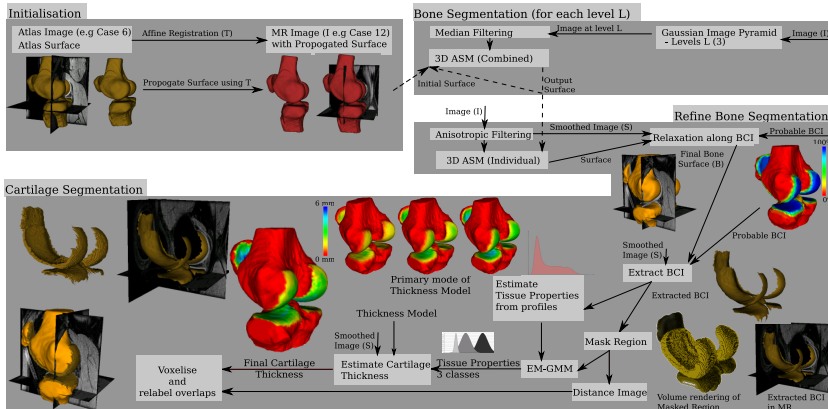
### 2.1 MR Acquisition

An inhomogeneous database of  $N = 20$  healthy volunteers who were not known to have knee problems were imaged using a FS SPGR MR sequence; 14 scanned at 1.5T (Tesla) and 6 scanned at 3T (Tesla). No demographic information was available, however no exclusion criteria based on age or gender was used. The following acquisition parameters were used in imaging: echo times - 5, 7 or 12 ms, repetition time - 60 ms, flip angle -  $40^\circ$ , Field of view 120 mm, matrix =  $512 \times 512$  or  $256 \times 256$ , 1.5 mm slice thickness, and 60, 64 or 70 sagittal slices. Each of the acquired images were manually segmented by an expert. Left knees were reflected across their axes and treated as right knees (Note: this is not currently handled automatically).

## 2.2 Segmentation System Overview

The major problem with segmenting the knee bones and cartilages from clinically relevant sequences (e.g. FS SPGR), is that both types of anatomy have poor or missing boundary interfaces. The bone is difficult to distinguish from fat, tendons and background, while the cartilage have problems with synovial fluid, meniscus and even ligaments and muscle tissue. In many applications statistical models [11] has proved to be useful at handling this type of problem. It should be noted that diseased knees are more difficult to segment, as osteophytes, lesions, intensity inhomogeneities and cracks are often found, however these are not present in our current database and our approach has not been validated on diseased knees.

This overall approach is illustrated in Figure 1 and is summarized as follows. A point distribution model represents the shape of the bones, whose variability across the database is modelled using 3D statistical shape models (SSM). A hybrid segmentation scheme based around 3D active shape models is used to segment the bones. Using the bone segmentation, the BCI is extracted and the local tissue properties of the bones and cartilages are estimated and represented using a Gaussian distribution. The initial tissue properties are then re-estimated using a 3 class expectation maximization based Gaussian mixture model performed in the local (mask) region around the BCI. This information, combined with constraints provided by a model (principal component analysis) of the thickness variation (observed in the database) is used to estimate the cartilage thickness at each point in the point distribution model. The outer cartilage surface is then extracted and this coupled bone-cartilage segmentation is voxelised. The SSM creation and bone segmentation was previously presented in [12] and this paper only provides a more detailed explanation of the cartilage segmentation.



**Fig. 1.** Flow diagram of the segmentation scheme used to automatically segment the bones, extract the BCIs and segment the cartilages. The surface rendering results presented are of case 12 with case 6 used as the atlas.

**Extraction of the BCI.** The BCI is the region of the bone that has articular cartilage above it. This region is determined automatically by first extracting the points on the bones which have a high probability of being on the BCI (determined from the training database). Using this initial BCI we then,

1. Estimate the cartilage tissue parameters which are modeled as a Gaussian distribution.
  - The distribution is estimated from samples extracted along a 1D profile  $g$ , normal to the surface at each point in BCI to the strongest negative gradient within 6.5 mm (supersampled at twice the in-plane resolution using cubic B-spline interpolation).
2. Consider all points who have two neighbors on the BCI (via triangulation).
  - If a point has at least 30% of the samples cartilage tissue then add point to current BCI.
3. Converged when the number of points on the BCI remains unchanged else Goto 2.

**Cartilage Segmentation.** The first stage of the segmentation process refines the estimate of the tissue properties used in the BCI extraction and generates a distance image from each of the bones. This is achieved by first creating a binary mask of voxels that are above and within 6.5 mm of the BCI. Inside this masked region an estimate of the tissue properties is obtained using an expectation maximization gaussian mixture model (3 classes, initialized using previous estimate of bones, cartilage and other (tissue between bone and cartilage intensities). From the tissue properties, a cartilage tissue probability image is generated (Gaussian based with values above the mean given a probability of 1). This is used as input to the cartilage segmentation algorithm, which operates as follows:  
Until range is 0.

1. For each point  $i$  on BCI find thickness  $t_i$ .
  - Assume that the position  $j$  along the profile  $g$  that maximizes the MR image gradient and internal probabilities ( $p$ ) ( $F(g_j, p_j) = \max(\frac{|\delta g_j|}{\max(|\delta g|)} + \frac{1}{k} \sum_k p_{j-k})$ ) corresponds to the correct outer cartilage edge.
2. Parameterize and reconstruct likely thickness using model of thickness variation (trained from normal patients).
3. Enforce BCI boundary constraints.
4. Set  $t_i$  to 0 for points that do not have an overall probability of having 50% cartilage tissue.
5. Decrease capture range by reducing search range around  $t_i$ .

After convergence the coupled bone-cartilage model is voxelised and the distance map is used to relabel overlapping voxels to their nearest cartilage interface.

## 2.3 Non-Rigid Registration

There are many ways of performing NRR, a popular method, due to it's general applicability, transparency and computational efficiency is the FFD approach first proposed by Rueckert [13]. In this paper, normalized mutual information was used with a total of 6 hierarchical levels (including the affine), each level decreased the spacing between the control points by half (20, 10, 5, 2.5 and 1 mm). The computational load was reduced by using a dilated mask to restrict the set of active control points.

## 2.4 Validation

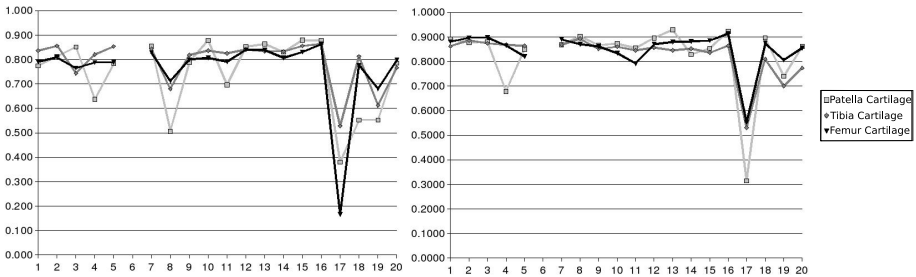
The validation of the segmentations are presented using three volume-based measures, sensitivity ( $= T_P / (T_P + F_N)$ ), specificity ( $= T_N / (T_N + F_P)$ ), DSC ( $= 2T_P / (2T_P + F_P + F_N)$ ) where  $T_P$  is true positive,  $T_N$  is true negative,  $F_P$

is false positive and  $F_N$  is false negative of the automatically obtained binary bone segmentation compared to the expert binary manual segmentation. The experiments were leave-one-out, with the robustness to initialization tested by segmenting each case 19 (N-1) times, each time using a different initialization (case as the atlas). Due to the high computational load, the NRR was only performed using two different atlases (case 6 and 15).

### 3 Results

As can be seen in Table 1 and Figure 2 our approach obtains much better sensitivity, specificity and DSC than NRR. The results obtained by our approach and NRR followed a similar trend across the cases (Figure 2) varying mainly depending on the image quality. A qualitative illustration of the difference in quality between our approach, manual segmentation and NRR is illustrated in Figure 3.

At low resolutions (cases 16 to 20) the results were slightly lower, however, the images were much noisier with much less contrast between the cartilages and surrounding tissue. This was especially true for cases 17 and 19 where the very low contrast resulted in none or very poor delineation (Figure 4) with few distinctive gradients found on the outer cartilage interfaces.

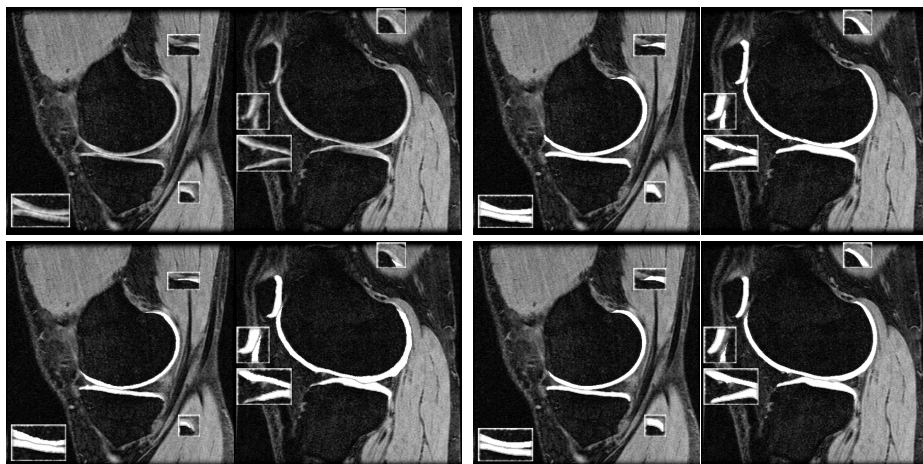


**Fig. 2.** The DSC for each case (case 6 as an atlas), showing that the *left* NRR after 1mm has inferior accuracy compared to *right* our approach. **Note 1:** The manual segmentations of the patellar cartilage for cases 4 and 17 appear to be incomplete.

Overall the primary cause of segmentation error was regions affected by partial voluming and other signal decreases. This occurred primarily around the edges and thin regions of the cartilages, it was also observed in regions of high curvature (in slice thickness direction). This problem could be reduced by using certain MR sequences and acquisition parameters. In a few cases small errors were also observed between the femoral and tibial cartilages (e.g Slice 48 in Figure 3). The approach was quite robust to initialization (atlas choice) with failures in cartilage segmentation only observed (with the exception of case 17) when the previous bone segmentation had failed (failure rate of bone 3.60% [12]).

**Table 1.** Mean (Standard deviation) (Median) of volume-based validations

Affine	Sensitivity		Specificity		DSC				
- (Patellar)	0.450	(0.163)	(0.502)	0.998	(0.001)	(0.998)	0.422	(0.164)	(0.491)
- (Tibial)	0.460	(0.170)	(0.491)	0.998	(0.001)	(0.998)	0.473	(0.166)	(0.519)
- (Femoral)	0.418	(0.143)	(0.474)	0.994	(0.002)	(0.994)	0.427	(0.138)	(0.478)
NRN (1 mm)	Sensitivity		Specificity		DSC				
- (Patellar)	0.803	(0.119)	(0.848)	0.999	(0.001)	(0.999)	0.732	(0.156)	(0.787)
- (Tibial)	0.781	(0.156)	(0.804)	0.999	(0.001)	(0.999)	0.785	(0.095)	(0.829)
- (Femoral)	0.795	(0.162)	(0.836)	0.997	(0.002)	(0.997)	0.758	(0.148)	(0.795)
Our approach	Sensitivity		Specificity		DSC				
- (Patellar)	0.821	(0.135)	(0.849)	1.000	(0.000)	(1.000)	0.833	(0.135)	(0.870)
- (Tibial)	0.829	(0.207)	(0.860)	0.999	(0.000)	(0.999)	0.826	(0.083)	(0.855)
- (Femoral)	0.837	(0.162)	(0.865)	0.999	(0.000)	(0.999)	0.848	(0.076)	(0.870)



**Fig. 3.** Overlaid segmentations (gray contour on patellar and tibial cartilages) for case 9 (case 15 as atlas, slices 16 and 48). *top left* MR *top right* Manual *bottom left* NRR (DSC=0.82,0.79,0.82) *bottom right* Our approach (DSC=0.87, 0.85, 0.86). **Note:** Areas of interest used 1.5 $\times$  zoom.

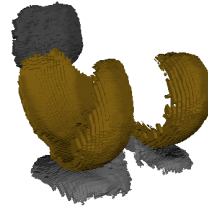
We further validated our approach using the data of the fourth subject used in [8]. This subject was scanned four times to evaluate the effect of partial voluming and repositioning, with the ground truth segmentation of each scan obtained using STAPLE from ten manual segmentations performed by two experts (each five times). As can be seen in Table 2, the values of sensitivity, specificity and DSC for the total cartilage (patellar, tibial and femoral) were similar, with the lower sensitivity and higher specificity of our approach indicating it slightly undersegments compared to the modified watershed algorithm. The primary advantage of our approach is that it does not require any user interaction and each



**Fig. 4.** Overlaid segmentation (gray contour on tibial cartilage) example of case 17 (case 6 used as atlas) using slice 20 *left* MR *middle* Manual *right* Our approach. **Note:** Poor delineation between cartilage interfaces cause our approach (and NRR) to fail.

**Table 2.** Average (from 5 segmentations) of the total cartilage (patellar, tibial and femoral) results obtained using our algorithm compared to the improved watershed approach of Grau [8]. Only the total cartilage was compared as Grau’s approach cannot obtain the individual cartilage segmentations that are necessary to perform statistically significant quantitative analysis

Scan	Improved Watershed			Our approach			$CV_{DSC}\%$
	Sens.	Spec.	DSC	Sens.	Spec.	DSC	
(1)	0.8965	0.9987	<b>0.8988</b>	0.8410	0.9993	<b>0.8897</b>	<b>0.06</b>
(2)	0.8649	0.9990	<b>0.8907</b>	0.8402	0.9994	<b>0.8898</b>	<b>0.08</b>
(3)	0.8763	0.9990	<b>0.8984</b>	0.8490	0.9992	<b>0.8902</b>	<b>0.1</b>
(4)	0.8905	0.9988	<b>0.8978</b>	0.8591	0.9992	<b>0.8959</b>	<b>0.09</b>



**Fig. 5.** Surface rendering of segmentation results obtained for case 9 (using case 15 as an atlas)

cartilage is segmented and labelled separately, which is essential for quantitative analysis.

The effectiveness of this approach for quantitative analysis was evaluated using two quantitative analysis measures (volume and thickness). The volume was estimated directly from which we found that the (manual, automatic) segmentations had an average volume of (4245, 3912), (6026, 6056) and (14703, 14463)  $mm^3$  and average absolute volume difference error of 8.29%, 4.94% and 5.56% for the patellar, tibial and femoral cartilages respectively (excluding case 17). The thickness was calculated from the whole BCI using an approach based on [14], from which obtained an average absolute thickness difference of (0.19, 0.33, 0.10  $mm$ ) for the (patellar, tibial, femoral) cartilage (excluding case 17).

## 4 Discussion and Conclusion

In this paper we have presented a segmentation scheme that automatically and accurately obtains cartilage segmentations from healthy volunteers. The

complete scheme is fully automatic and has a reasonable computation time (slightly over an hour). Each cartilage is obtained as a separate object, which is essential for performing statistically significant quantitative analysis. We also found that the results did not vary significantly due to repositioning and partial voluming effects, which is critical as these effects are impossible to avoid. Using this scheme for quantitative analysis obtained an average absolute volume difference error of only 8.29%, 4.94% and 5.56%.

Future work is focused on further improving the segmentation results by increasing the size of the training database and including additional information, including localized tissue and texture models and explicit curvature constraints. This is primarily aimed at improving the segmentation accuracy at the poorly delineated cartilage-cartilage interfaces and towards the thinner cartilage boundaries. Future research will be focused on validating this scheme on the more complex problem of diseased knees.

## Acknowledgment

The authors wish to thank Andrea U. J. Mewes and Johannes Pauser for help in acquiring and interactively segmenting the MR scans. This investigation was supported in part by NIH grants R21 MH067054 and R01 RR021885, a research grant from CIMIT and grant RG 3478A2/2 from the NMSS. The authors would also like to thank Mark Holden for help in using the VTK CISG Segmentation Propagation Tool [15] to perform the non-rigid registration.

## References

1. Eckstein, F., et al.: Magnetic resonance imaging (MRI) of cartilage in knee osteoarthritis (OA): morphological assessment. *Osteoarthritis and Cartilage* 14, 46–75 (2006)
2. Cicuttini, F., et al.: Comparison of conventional standing knee radiographs and magnetic resonance imaging in assessing progression of tibiofemoral joint osteoarthritis. *Osteoarthritis Cartilage* 13(8), 722–727 (2005)
3. Waterton, J., et al.: Diurnal variation in the femoral articular cartilage of the knee in young adult humans. *Magnetic Resonance in Medicine* 43, 126–132 (2000)
4. Stammberger, T., et al.: Determination of 3D cartilage thickness data from MR imaging: computational method and reproducibility in the living. *Magnetic Resonance in Medicine* 41(3), 529–536 (1999)
5. Gougoutas, A., et al.: Cartilage volume quantification via live wire segmentation. *Academic Radiology* 11(12), 1389–1395 (2004)
6. Yoshioka, H., et al.: Articular cartilage of knee: Normal patterns at MR imaging that mimic disease in healthy subjects and patients with osteoarthritis. *Radiology* 231, 31–38 (2004)
7. Lang, P., et al.: MR Imaging of Articular Cartilage: Current State and Recent Developments. *Radiologic Clinics of North America* 43(4), 629–639 (2005)
8. Grau, V., et al.: Improved watershed transform for medical image segmentation using prior information. *IEEE Trans. Medical Imaging* 23(4), 447–458 (2004)



9. Folkesson, J., et al.: Segmenting articular cartilage automatically using a voxel classification approach. *IEEE Trans. Medical Imaging* 26(1), 106–115 (2007)
10. Li, K., et al.: Simultaneous segmentation of multiple closed surfaces using optimal graph searching. In: Christensen, G.E., Sonka, M. (eds.) *IPMI 2005*. LNCS, vol. 3565, pp. 406–417. Springer, Heidelberg (2005)
11. Cootes, T., et al.: Active shape models - their training and application. *Computer Vision and Image Understanding* 61(1), 38–59 (1995)
12. Fripp, J., et al.: 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)
13. Rueckert, D., et al.: Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images. *IEEE Transactions on Medical Imaging* 18(8), 712–721 (1999)
14. Jones, S., et al.: Three-dimensional mapping of cortical thickness using Laplace's equation. *HBM* 11, 12–32 (2000)
15. Hartkens, T.: VTK CISG Registration Toolkit (2006), Software available at <http://freshmeat.net/projects/vtkcisg/>