Segmenting Tibia and Femur from Knee X-ray Images

Matthias Seise[1]*, Stephen J. McKenna[1], Ian W. Ricketts[1] and Carlos A. Wigderowitz[2]

[1]Division of Applied Computing, University of Dundee, Dundee, Scotland, DD1 4HN
[2]Orthopaedic and Trauma Surgery, Ninewells Hospital and Medical School, Dundee, Scotland, DD1 9SY

Abstract. Automatic segmentation of bone contours in knee x-ray images is investigated as a step towards reliable, quantitative radiographic analysis of osteoarthritis for diagnosis and assessment of progression. A double contour active shape model is proposed in order to simultaneously segment anterior and posterior contours of the tibial plateaux. Several features are compared for modelling local appearance. Point-to-contour segmentation errors are reported for both femoral and tibial contours.

1 Introduction

Osteoarthritis (OA) is the most common joint disease and the most common cause of disability in older people [1], resulting in significant economic costs for society. It is characterised by an imbalance in the synthesis and degeneration of articular cartilage. In OA of the knee, cartilage covering the tibial plateaux and femoral condyles (see Figure 1) is typically destroyed. Two-dimensional x-ray imaging is the most widely used modality for assessing progression of OA. Since cartilage is not apparent in x-ray images, the primary radiographic sign used is the apparent space between the femoral condyles and the tibial plateaux. This space tends to shrink as the cartilage is destroyed. Further radiographic signs are sclerotic bone, cysts and osteophytes (bone spurs due to growth of normal bone along weight-bearing regions of the tibia and femur). Although automated measurement of joint space width parameters has been attempted [2], reliable and objective methods for quantitative analysis of OA progression based on radiographic signs are not available. This paper deals with automated segmentation of contours of the tibia and femur as an important step towards this goal.

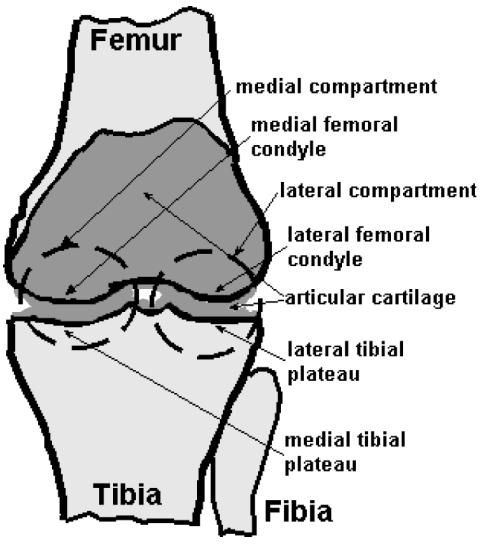
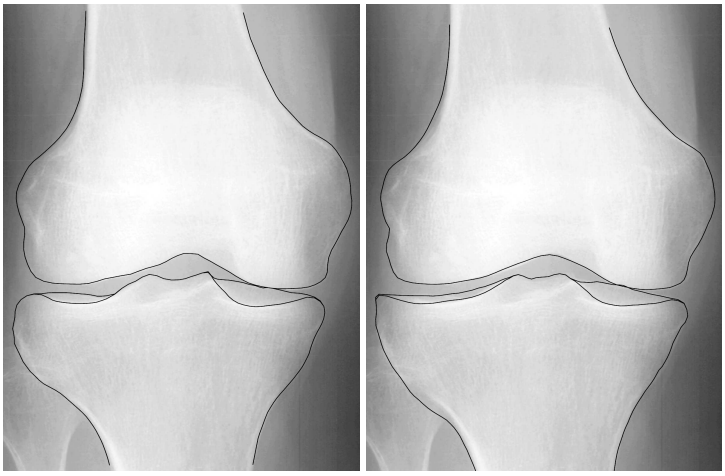


Figure 1. Anatomy of the knee



(a) (b)

Figure 2. A standard clinical x-ray, contrast enhanced for visualisation purposes only. (a) A manual annotation of femoral and tibial contours. (b) A segmentation result using DCASM.

Figure 2(a) shows a manual annotation of the femoral and tibial contours of interest in a clinical anteroposterior (AP) x-ray image. In this example, concavities in the tibial plateaux result in distinct image contours corresponding to the anterior and posterior rims of the plateaux. It is very difficult to determine which contour is which on the basis of the AP radiograph. The contours are therefore referred to as the *inner* contour and the *outer* contour in a 2D sense. These double contours are not always present on both the lateral and medial plateaux. Furthermore, the contour bifurcation points vary quite widely between example images. An extension to the active shape model (ASM) [3, 4], referred to as the double contour active shape model (DCASM), was developed for modelling and segmentation of such contours. The remainder of the paper describes this model and its application to a set of standard clinical radiographs. Figure 2(b) shows an example segmentation result obtained.

*Corresponding author: M. Seise. email: mseise@computing.dundee.ac.uk

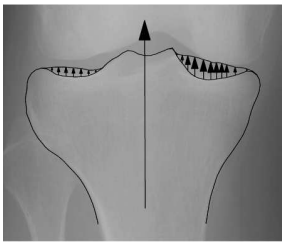


Figure 3. Dominant axis

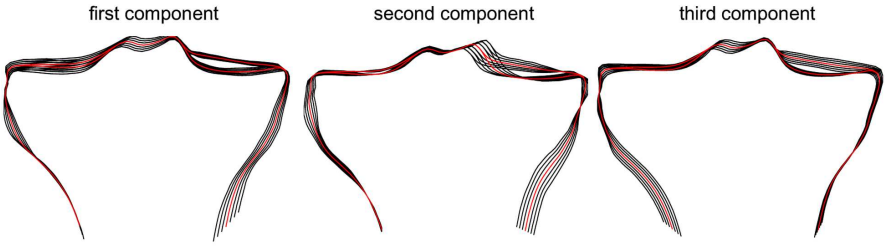


Figure 4. Modes of variation of tibia for the largest 3 eigenvalues

2 Double Contour Active Shape Models

In common with the standard ASM procedure, a set of training examples is first brought into alignment using Procrustes analysis based on landmark points extracted along manually annotated contours. The landmarks and endpoints were determined using the minimum description length (MDL) approach of Davies *et al.* [5] with use of curvature [6]. Contours with loops such as in Figure 2(a) are treated as double contours. The Procrustes alignment and MDL landmarking process is applied only to the inner contour which is treated as a reference contour. Each (aligned) reference contour is thus described by N landmark points $\{(x_n, y_n)\}_{n=1}^{N}$.

Instead of treating the outer contour independently, the displacement required to move each of the N landmarks onto the outer contour is determined. In the implementation described here, these displacements are measured along a shared dominant axis direction α which is determined so that this representation is well-defined (see Figure 3). The s^{th} training shape is then represented as the $3N$-vector $\mathbf{x}_s = (x_1, y_1, \delta_1, \ldots, x_N, y_N, \delta_N)^\top$. The sample mean and the covariance matrix are computed using Equations (1).

$$\overline{\mathbf{x}} = \frac{1}{S}\sum_{s=1}^{S}\mathbf{x}_s \qquad\qquad \mathbf{C} = \frac{1}{S-1}\sum_{s=1}^{S}(\mathbf{x}_s - \overline{\mathbf{x}})(\mathbf{x}_s - \overline{\mathbf{x}})^\top \tag{1}$$

Let $\mathbf{\Phi} = (\phi_1|\phi_2|\cdots|\phi_D)$ denote the matrix whose columns are the D eigenvectors corresponding to the D largest eigenvalues $\lambda_1, \ldots, \lambda_D$ of $\mathbf{C}$. The number of eigenvectors to retain is calculated as the smallest D such that 95% of the total variance is explained. Any example of the training set, $\mathbf{x}_s$, can be approximated by

$$\mathbf{x}_s \approx \overline{\mathbf{x}} + \mathbf{\Phi}\mathbf{b}_s \tag{2}$$

where $\mathbf{b}_s$ is the D-dimensional model parameter vector, computed by

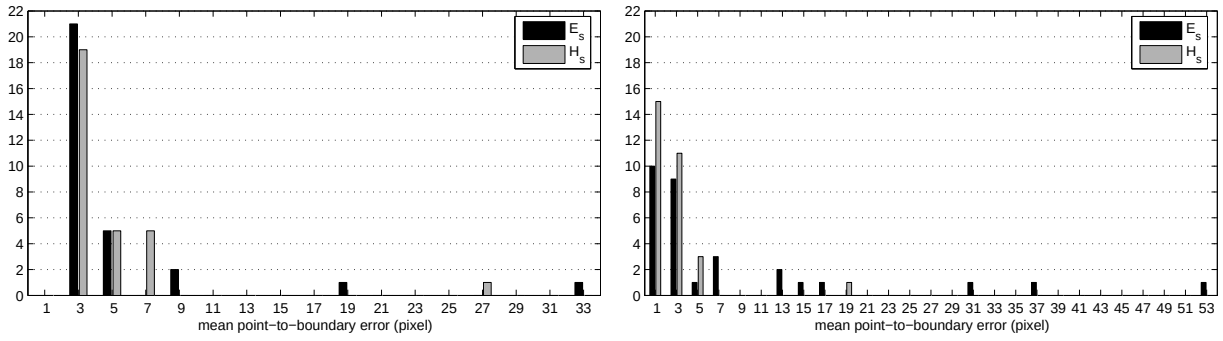
$$\mathbf{b}_s = \mathbf{\Phi}^\top(\mathbf{x}_s - \overline{\mathbf{x}}) \tag{3}$$

It is standard practice when building an ASM to model each landmark's local appearance using a 1-D profile centred on and orthogonal to the contour. The appearance is sampled on the contour and at K points to either side of the contour, giving a profile of length $2K + 1$. This approach is used in the DCASM for modelling the local appearance only at those landmark points with zero displacement for all training examples (i.e. when $\delta_n = 0$ for all training examples). For the remaining landmarks, profiles in the direction of the dominant axis, α, and centred on both the inner and outer contours are used. The dominant axis direction is used because during each search the landmark points are constrained to move in this direction. Adopting another direction would necessitate a complicated and numerically unstable recalculation of the displacement in each search step. The dominant axis direction is often similar to the inner contour normal direction so the resulting appearance models are also similar.

A popular choice of appearance feature for ASM profiles is the normalised first-order derivative (normalised gradient) [4]. However, Behiels *et al.* [7] reported significantly better segmentation of bones in x-ray images using alternative features. Therefore, several different features were compared here, namely raw intensity, unnormalised gradient, normalised intensity, normalised gradient, scaled intensity and scaled gradient[1].

When the number of training examples is limited, appearance models learned separately for each landmark can become unreliable. A windowing method is therefore adopted in which training profiles from nearby landmarks

[1]For $\mathbf{g} = (g_1, \ldots, g_M)$, the normalised vector is $\hat{\mathbf{g}} = \left(\frac{g_1}{\sum_{m=1}^{M} g_m}, \ldots, \frac{g_M}{\sum_{m=1}^{M} g_m}\right)$ and the scaled vector is $\tilde{\mathbf{g}} = \left(\frac{g_1 - \min \mathbf{g}}{\max \mathbf{g} - \min \mathbf{g}}, \ldots, \frac{g_M - \min \mathbf{g}}{\max \mathbf{g} - \min \mathbf{g}}\right)$



(a) (b)

Figure 5. Histograms (bin width $= 2$) of the segmentation errors E_s and H_s when using scaled gradient and $W = 6$. (a) Tibia. (b) Femur.

are pooled in order to estimate the appearance model. More specifically, for each landmark, profiles from the W adjacent landmarks to its left and the W landmarks to its right on the contour are used in addition to profiles at the landmark itself in order to estimate a mean profile and covariance matrix. This windowing is used for the single contour landmarks as well as the double contour landmarks.

A standard multi-resolution ASM search [4] is used along with the modifications needed to accommodate the double contour model. It should be noted that a standard ASM is recovered as a special case of the DCASM when all the displacements are set to zero.

3 Experimental Evaluation

The methods described above were evaluated on a data set of 30 standard clinical x-rays of normal knees. The image shown in Figure 2(a) is 510×740 pixels which gives an idea of the resolution used. Images of left knees were mirrored so that they appeared as right knees. All images were manually annotated and leave-one-out validation was used for evaluation. One highly approximate, manual initialisation was provided for each image.

Segmentation accuracy is reported here in terms of the mean point-to-boundary error, E_s, which is the average distance in pixels from the obtained landmark positions to the annotated contour (taken as ground truth) for the s^{th} example. This error averaged over the 30 test examples is denoted $\overline{E}$. For the application of measuring the joint space, segmentation of the tibial plateaux and the femoral condyles is of particular importance. Therefore, the mean error along these sections of the contours is also reported and is denoted H_s.

A typical overall segmentation result is shown in Figure 2(b). The first three modes of variation obtained for the tibia are illustrated in Figure 4. Figure 5(a) plots the distribution of segmentation errors obtained for the tibia using $W = 6$ and scaled gradient profiles. Note that three examples had very poor segmentations and two of these are shown in Figure 7. The median of the errors E_s for the tibia test examples was 3.3 pixels ($\min = 2.1$, $\max = 34.0$). The median of the error H_s for the tibial plateaux was 3.2 pixels ($\min = 2.1$, $\max = 26.4$).

Figure 5(b) shows the segmentation errors obtained for the femur using scaled gradient profiles and the windowing parameter $W = 6$. The median of the errors E_s was 2.3 pixels ($\min = 1.4$, $\max = 52.3$). The median of the errors H_s for the femoral condyles was 2.0 pixels ($\min = 1.1$, $\max = 18.4$). The worst two results are shown in Figure 7.

Figure 6 plots the effect on the tibia segmentation accuracy of varying the windowing parameter and the choice of appearance feature. The windowing parameter W did not have a large impact. However, the different features resulted in quite different segmentation accuracies. Scaled gradient was best and normalised intensity was worst for both tibia and femur.

4 Discussion and Conclusions

This paper has shown that the proposed double contour active shape model can be used to segment contours of the tibia and femur in knee x-rays. This is the first published use, to the authors' knowledge, of statistical shape

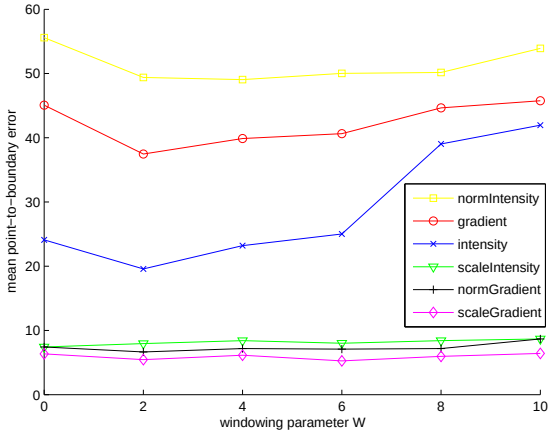


Figure 6. The effect of profile feature type and window size on the mean point-to-contour errors, $\overline{E}$, when segmenting the tibia.

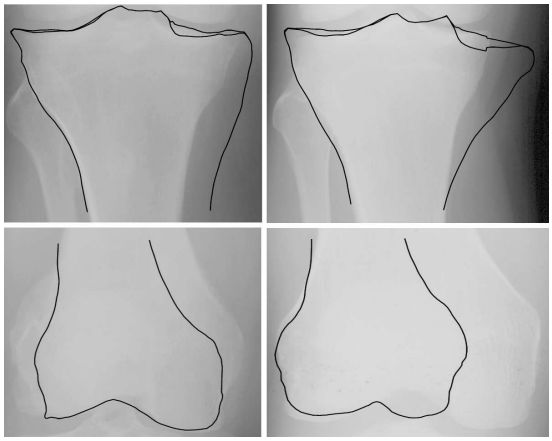


Figure 7. Segmentation failures using scaled gradient and $W = 6$. Top: tibia. Bottom: femur.

models to segment tibia and femur in plane radiographs of the knee joint. The initial results are promising. In particular, the errors along the tibial plateaux and the femoral condyles are in many cases smaller than 4 pixels for the tibia (19 cases) and 3 pixels for the femur (24 cases). Subsequent work is needed to determine whether this is sufficient for useful joint space measurement.

Tibia segmentation failures appeared to be due to structured tissue outside the bone which is clearly visible in the upper right image of Figure 7. Since the shape converges to this false contour in low resolutions it cannot find the true contour at higher resolutions. Therefore, it converges to a false contour. The correct segmentation might be recovered by using a different initialisation or using longer profiles, for example.

The finding that scaled gradient was best and that normalised intensity was worst stands in contrast to Ref. [7] in which scaled intensity seemed to perform best for segmenting the *upper* end of the femur. This demonstrates that even for similar applications, the optimal appearance models can be different and difficult to find. Future work will include development of improved appearance models and search methods for this application. Since the appearance distribution is non-Gaussian, better quality of fit measurements than the Mahalanobis distance should be possible. Promising initial results are being obtained using k-nearest neighbour weighted local regression.

Acknowledgments

The authors acknowledge the EPSRC who partially funded M. Seise, H. H. Thodberg who provided the code for the MDL method and Dr B. Oliver who contributed the normal knee image data.

References

1. T. E. McAlindon, C. Cooper, J. R. Kirwan et al. “Knee pain and disability in the community.” *Br J Rheumatol* **31(3)**, pp. 189 – 92, Mar 1992.
2. J. Duryea, S. Zaim & H. K. Genant. “New radiographic-based surrogate outcome measures for osteoarthritis of the knee.” *Osteoarthritis Cartilage* **11(2)**, pp. 102 – 110, Feb 2003.
3. T. F. Cootes, C. J. Taylor, D. H. Cooper et al. “Active shape models — their training and application.” *Computer Vision and Image Understanding* **61(1)**, pp. 38 – 59, Jan 1995.
4. T. F. Cootes & C. J. Taylor. “Statistical models of appearance for computer vision.” Technical report, Uni. of Manchester, 2004.
5. R. H. Davies, C. Twining, T. F. Cootes et al. “A minimum description length approach to statistical shape modelling.” *IEEE Transactions on Medical Imaging* **21(5)**, pp. 525 – 537, May 2002.
6. H. H. Thodberg & H. Olafsdottir. “Adding curvature to minimum description length shape models.” In R. Harvey & A. Bangham (editors), *British Machine Vision Conference*, volume 2, pp. 251 – 260. Sep 2003.
7. G. Behiels, F. Maes, D. Vandermeulen et al. “Evaluation of image features and search strategies for segmentation of bone structures in radiographs using active shape models.” *Medical Image Analysis* **6(1)**, pp. 47–62, 2002.