Original Research Paper

Comparison of Improved Semi-Automated Segmentation Technique with Manual Segmentation: Data from the Osteoarthritis Initiative

1Hong Seng Gan and 2Khairil Amir Sayuti
1Universiti Kuala Lumpur, British Malaysian Institute, Medical Engineering Technology Section, 53100 Gombak, Selangor, Malaysia
2Department of Pathology, School of Medical Science, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia

Abstract: Manual segmentation is the standard procedure in osteoarthritis study. However, this method is infamous for being excessive, time consuming and exhaustive. In this study, we overcame the problem of excessive expert interaction reported in manual segmentation by developing a semi-automated random walks technique with computer-aided labelling system. To minimize expert interaction, non-cartilage seeds were generated by using computer. Then, random walks algorithm would segment knee cartilage based on cartilage seeds and non-cartilage seeds. Finally, segmentation results were revised and refined accordingly. A total of 15 normal images and 10 osteoarthritic images were used in this study. In term of efficiency, we have reduced the processing time to segment normal cartilage by 47.5% (93±21s; P = 0.0000019) for observer 1 and 44% (61±8s; P = 3.52×10−5) for observer 2. We also reduced the processing time to segment diseased cartilage by 48.1% (56±16s; P = 0.00014) for observer 1 and 30.3% (62±14s; P = 0.0070) for observer 2. Besides, the proposed technique have produced good reproducibility in both normal (0.83±0.028 for observer 1 and 0.80±0.040 for observer 2) and diseased (0.80±0.060 for observer 1 and 0.82±0.043 for observer 2) cartilage segmentations. In conclusion, the combination of computer generated seeds and user-friendly random walks method have reduced the amount of expert interaction to necessary level without compromising the accuracy of results.

Keywords: Semi-Automated Segmentation, MR Image, Seeds, Osteoarthritis, Knee Cartilage

Introduction

Osteoarthritis (OA) is one of the most complicated joint disease ever studied by mankind (Brandt et al., 2009; Valdes and Spector, 2009; Shane Anderson and Loeser, 2010; Linn et al., 2012; Sokolove and Lepus, 2013). Without effective cure for this diarthrosis disease until now, the implications associated with OA are destructive from social (Losina et al., 2011) and economic (Kotlarz et al., 2009) perspectives. Therefore, researchers are keen on Developing Disease Modifying Osteoarthritis Drug (DMOAD) to halt the prevalence of osteoarthritis. To develop DMOAD, potential imaging biomarkers based on Magnetic Resonance (MR) imaging have been studied quantitatively (Wang et al., 2012).

Intuitively, quantification of OA progression requires cartilage to be extracted from the MR image. Manual segmentation is the standard procedure but this practice is excessive, time-consuming and exhaustive. Subsequently, semi-automated and fully automated methods have been proposed to replace manual segmentation. Semi-automated methods such as active contour (Stammberger et al., 1999), graph cuts (Bae et al., 2009) and shortest paths (Gougoutas et al., 2004) require expert initialization and let algorithms to separate the cartilage from MR image. In the case of osteoarthritis study, failure of existing algorithms to cope with complicated anatomical cartilage geometry leads to high level of expert intervention. Meanwhile, automatic
methods such as Approximate Nearest Neighbours (ANN) (Folkesson et al., 2007), active model (Solloway et al., 1997) and atlas method (Carballido-Gamio and Majumdar, 2011) rely heavily on artificial training to initialize and perform cartilage segmentation. Hence, development of fully-automated segmentation techniques are often complex and depend on accuracy of training dataset.

Since the development of reliable automated techniques remain fledging and semi-automated techniques still demand high degree of expert interaction, we developed a semi-automated segmentation model supported by computer-generated non-cartilage seeds. Consequently, expert role will be reduced to placement of cartilage seeds. In the following sections, we give step-by-step explanation about the development of segmentation model in Material and Methods section and types of evaluation performed in Result section. Next, we discuss about the result at Discussion of the Result and lastly, conclude our works.

Materials and Methods

Image Dataset

All images used in this study were Dual Echo Steady State (DESS) MR knee images with water excitation (we) obtained from the Osteoarthritis Initiative (OAI) dataset. These images, with approval from Institutional Review Board, were acquired by using 3.0 Tesla (T) MRI scanner (Siemens Magnetom Trio, Erlangen, Germany) with quadrature transmit-receive knee coil (USA Instruments, Aurora, OH) (Eckstein et al., 2006). The DESS images have section thickness of 0.7 mm and an in-plane resolution 0.365×0.456 mm² (field of view = 140×140 mm, flip angle = 25°, TR/TE = 16.3/4.7 m sec, matrix size = 384×384 mm, bandwidth = 185 Hz/pixel) (Peterfy et al., 2008). More information related to the Osteoarthritis Initiative (OAI) dataset can be found at http://oai.epi-ucsf.org/datarelease/About.asp.

Segmentation Model

The proposed method was developed by using MATLAB R2014a (Math works, Natick, MA) on laptop equipped with Corei7-4700HQ@2.50 GHz processor and 8.00 GB RAM. Procedures to conduct manual and improved semi-automated knee cartilage segmentation were described in Fig. 1. Accordingly, MR image is loaded into an interactive tool (Hong-Seng et al., 2014) with manual and semi-automated segmentation modes. Different from conventional computer-aided segmentation, the improved model is supported by computer-generated non-cartilage seeds and manual cartilage seeds. Both types of seeds will provide pertinent feature information to initiate random walks algorithm. Meanwhile, manual segmentation is executed through boundary delineation function where experts are required to draw the knee cartilage boundary meticulously and refine the result until satisfaction level. Lastly, segmentation results from both techniques are studied quantitatively using similarity index.

---

Fig. 1. Execution of manual segmentation method and improved semi-automated segmentation method
Automated Non-cartilage Seeds Generation

Computer-generated seeds (Gan et al., 2014a) are developed for non-cartilage compound based on two consistent patterns. First, non-cartilage seeds will always enclose the cartilage. Second, position of cartilage is always located at the centre of MR image. At initial stage, MR image is divided into \( k = 200 \) homogenous and orderly structured super pixels. Then, Fuzzy C-Means (FCM) algorithm is implemented inside each super pixel. The objective function of FCM is shown in (1):

\[
J_m = \sum_{i=1}^{N} \sum_{j=1}^{CC} \mu_{ij}^m \| x_i - cc_j \|^2
\]

Where:
- \( N \) indicates total number of pixels in non-cartilage super pixels
- \( CC \) - The number of cluster inside each super pixels
- \( \mu_{ij}^m \) - Governs the percentage of belonging of a given point to each cluster
- \( m \) - The degree of membership
- \( x_i \) - The pixel location
- \( cc_j \) - The cluster center

As a result, clusters will be created inside each super pixel. In this context, we treat the cluster centroids as potential seeds and the accuracy of cluster formation does not affect the seed generation model. The location of seed, however, does not necessarily reflect actual location. Hence, the seeds are approximated to image pixel with minimum distance. The definition is given in (2):

\[
dist(ij) = \min \left( \sum (x_i - cc_j) \right)^2
\]

Manual Labelling by Expert

The objective of interactive labelling is to identify different types of cartilage components (i.e., femoral, tibial and patellar) using expert knowledge. Because femoral, tibial and patellar cartilage seeds are developed independently, they have different colours to avoid confusion. The seeds are inserted in the form of curvilinear lines through an intuitive user interface (Hong-Seng et al., 2014). Despite the seeds are placed within the object of interest, shapes of the seeds do not imply actual result but serve as hard constraint to guide the segmentation algorithm. Since non-cartilage component will be covered by computer-generated seeds, expert can focus on labelling cartilage compound only. Fig. 2a demonstrated the complex labelling effort consisting of non-cartilage and cartilage labelling required by conventional random walks while Fig. 2b shows the simplified labelling procedures after the improved semi-automated method is implemented to handle non-cartilage labelling. As a result, expert only needs to insert femoral, tibial and patellar cartilage labels.

Implementation of Random Walks

Random walks is a graph based segmentation method. As such, MR image is treated as a graph structure \( G = (V, E) \) where \( V \) is the node/pixel and \( E \) is the edge between two pixels. To differentiate two objects, the algorithm measures the probability that a pixel at specified location belongs to different types of labels by solving the system of linear equation as shown in (3):

\[
L_U X = -B^T M
\]

Where:
- \( L_U \) - A Laplacian matrix storing degree and weight information of unlabelled seeds
- \( X \) - Matrix containing the probabilities of unlabelled seeds
- \( B \) - A matrix decomposition of \( L \) and \( M \) is a matrix containing the probabilities of labelled seeds.

Probability maps is computed for each type of label \( j \) where \( j = 1, 2, 3, 4 \) or 5. The sum of probability at any pixel location \( i \) equals to unity as shown in (4):

\[
\sum_{j=1}^{5} x'_i = 1, \forall v_i \in V
\]

Fig. 3a-d shows the probability map produced by femoral, tibial, patellar and non-cartilage labels respectively using improved semi-automated method where darker intensity indicates lower probability (Gan et al., 2014b). Eventually, the pixel will be assigned to label with highest probability.

Refinement of Seeds

Sometimes, experts may find the seeds being placed wrongly during labelling or the result deemed erroneous after segmentation. Therefore, refinement is carried out by revising the seed placement through deletion or addition of seeds and then re-generate the result. For example, Fig. 4a demonstrated the initial scribble of different labels. Nevertheless, in Fig. 4b, femoral cartilage was found to be over segmented as indicated by blue arrow. To correct the result, tibial label can be removed as indicated by blue arrow in Fig. 4c and d and redraw to cover previously under segmented region in Fig. 4e. Eventually, final result was exhibited in Fig. 4f.
Fig. 2. Comparison between redundant traditional labelling procedures used by random walks method and simplified labelling procedures introduced by the improved method

Fig. 3. Computation of probability map for each label

Fig. 4. Demonstration of refinement procedures when erroneous result is found
Results

Evaluation Metric

In this experiment, 15 normal images and 10 diseased images were chosen randomly from the OAI dataset. Two experts (K.A.S and A.H.A.K) independently performed the segmentation of knee cartilage using the improved semi-automated method and manual method. Both experts were blinded throughout the experiment. Performances were assessed using efficiency and reproducibility tests. For semi-automated method, efficiency was defined as the total processing time needed for an expert to place seeds, compute segmentation and refine results until satisfaction. For manual method, efficiency was defined as the total time required for an expert to delineate the boundary of different types of cartilage. The reproducibility of manual and semi-automated method was measured by using Dice similarity coefficient, sensitivity and specificity. Dice similarity coefficient measures the degree of agreement between two sets of results, sensitivity measures the correct classification of cartilage pixels and specificity measures the correct classification of non-cartilage pixels. Formulations of Dice’s coefficient, sensitivity and specificity are indicated in (5), (6) and (7) respectively:

\[
\text{Dice} = \frac{2TP}{(FP + TP) + (TP + FN)} \tag{5}
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \tag{7}
\]

where, TP is true positive, TN is true negative, FP is false positive and FN is false negative. A Dice value of 0 indicates no overlap while a Dice value of 1 indicates perfect agreement. Sensitivity and specificity have maximum value of 1 and minimum value of 0.

Statistical Analysis

All statistical analyses were performed by using SPSS (version 21; SAS Institute, Cary, NC) where α<0.05 indicated significant difference. The efficiency assessment was further divided into normal and diseased categories. In each category, time needed by both observers to segment normal cartilage using manual method and the improved semi-automated method were compared. Then, Pearson’s coefficient was computed to evaluate the correlation between normal cartilage data of observer 1 and observer 2. Pair t test was computed to study the significant difference between manual and the improved semi-automated method.

Reproducibility was examined from intra-observer perspectives by superimposing segmentation result generated by manual method on results generated by semi-automated methods. The analysis was carried out on global cartilage level and compartmentalized cartilage level. At each level, the analysis was further specified into normal (Kellgren-Lawrence, KL grade = 0) and diseased (KL = 1-4) cases. Similarly, pair t test was performed to examine the significant difference between two sets of data.

Segmentation Efficiency

Table 1 summarized the mean processing time required by manual and semi-automated method. In normal cartilage category, greater efficiency was achieved by using the improved semi-automated method (MS: 177±42s against PM: 93±21s; \(P = 0.0000019\)) for observer 1. Meanwhile, observer 2 consumed 109±29s to segment normal cartilage using manual method but the efficiency was improved significantly by using the proposed method (61±8s; \(P = 3.52 \times 10^{-5}\)). Pearson’s coefficients between observer 1 and 2’s results were 0.15 and 0.29 by using manual method and the proposed method respectively, indicating that there was no significant correlation between both sets of data.

In diseased cartilage category, observer 1 spend 108±22s to segment the cartilage when using manual method. The processing time was reduced significantly by using the proposed method (56±16s; \(P = 0.00014\)). Meanwhile, observer 2 spend 89±20s to segment diseased cartilage but the segmentation process was significantly improvised by using the proposed method (62±14s; \(P = 0.0070\)). Similarly, no significant correlation was found inefficiency data generated by observer 1 and 2 given that Pearson’s coefficients between observer 1 and 2’s data were 0.22 and 0.30 for manual method and the proposed method.

Segmentation Reproducibility

Table 2 summarized the mean reproducibility generated by global cartilage. The proposed method, for instance, has demonstrated good level of agreement with manual segmentation results. In normal cartilage section, the proposed method has reported reproducibility of 0.83±0.028 for observer 1 and 0.80±0.060 for observer 2. In diseased cartilage section, the proposed method has reported mean reproducibility of 0.80±0.060 for observer 1 and 0.82±0.043 for observer 2.

Table 3 demonstrated mean reproducibility generated by compartmentalized cartilage data. In normal cartilage category, observer 1 produced reproducibility of 0.82±0.038, 0.82±0.048 and 0.81±0.069 when segmenting normal femoral, tibal and patellar cartilage respectively by applying the improved semi-automated method while observer 2 generated reproducibility of 0.82±0.039, 0.75±0.054 and 0.78±0.10 when segmenting normal femoral, tibal and patellar cartilage.
Table 1. Processing time (in second) recorded by Manual Segmentation (MS) and the Proposed Method (PM) (standard deviation)

<table>
<thead>
<tr>
<th>KL grade</th>
<th>Observer 1</th>
<th>Observer 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MS</td>
<td>PM</td>
</tr>
<tr>
<td>0</td>
<td>177s (42s)</td>
<td>93s (21s)</td>
</tr>
<tr>
<td>1-4</td>
<td>108s (22s)</td>
<td>56s (16s)</td>
</tr>
</tbody>
</table>

Table 2. Measurement of reproducibility (standard deviation) for global cartilage

<table>
<thead>
<tr>
<th>DSC</th>
<th>Sens</th>
<th>Spec</th>
<th>DSC</th>
<th>Sens</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>KL = 0</td>
<td>0.83 (0.028)</td>
<td>0.84 (0.037)</td>
<td>0.996 (0.0014)</td>
<td>0.82 (0.040)</td>
<td>0.87 (0.051)</td>
</tr>
<tr>
<td>KL = 1 - 4</td>
<td>0.80 (0.060)</td>
<td>0.86 (0.044)</td>
<td>0.994 (0.0019)</td>
<td>0.82 (0.043)</td>
<td>0.85 (0.049)</td>
</tr>
</tbody>
</table>

Table 3. Measurement of reproducibility (standard deviation) for compartmentalized cartilage

<table>
<thead>
<tr>
<th>Label</th>
<th>DSC</th>
<th>Sens</th>
<th>Spec</th>
<th>DSC</th>
<th>Sens</th>
<th>Spec</th>
</tr>
</thead>
<tbody>
<tr>
<td>KL = 0</td>
<td>i = 1</td>
<td>0.82 (0.038)</td>
<td>0.86 (0.053)</td>
<td>0.987 (0.0038)</td>
<td>0.82 (0.039)</td>
<td>0.84 (0.050)</td>
</tr>
<tr>
<td></td>
<td>i = 2</td>
<td>0.82 (0.048)</td>
<td>0.84 (0.053)</td>
<td>0.999 (0.00036)</td>
<td>0.75 (0.065)</td>
<td>0.82 (0.071)</td>
</tr>
<tr>
<td></td>
<td>i = 3</td>
<td>0.81 (0.069)</td>
<td>0.84 (0.10)</td>
<td>0.999 (0.00043)</td>
<td>0.78 (0.10)</td>
<td>0.87 (0.16)</td>
</tr>
<tr>
<td>KL = 1-4</td>
<td>i = 1</td>
<td>0.80 (0.077)</td>
<td>0.86 (0.048)</td>
<td>0.996 (0.0018)</td>
<td>0.81 (0.047)</td>
<td>0.81 (0.053)</td>
</tr>
<tr>
<td></td>
<td>i = 2</td>
<td>0.76 (0.069)</td>
<td>0.82 (0.10)</td>
<td>0.998 (0.00079)</td>
<td>0.72 (0.081)</td>
<td>0.81 (0.051)</td>
</tr>
</tbody>
</table>
|        | i = 3    | 0.75 (0.048) | 0.78 (0.13)  | 0.999 (0.00047) | 0.80 (0.085) | 0.88 (0.088) | 0.999 (0.0039) 

In diseased cartilage category, observer 1 has generated reproducibility of 0.80±0.077, 0.76±0.069 and 0.75±0.048 when segmenting diseased femoral, tibial and patellar cartilage respectively. Observer 2 has generated reproducibility of 0.81±0.047, 0.72±0.081 and 0.80±0.085 when segmenting diseased femoral, tibial and patellar cartilage respectively.

Discussion of the Results

Compared to manual method and automatic method, semi-automated method represents the most practical segmentation approach to segment cartilage. Direct incorporation of expert supervision overcomes anatomical irregularities and pathological features exhibited by knee cartilage. However, we found out that redundancy in existing labelling procedure is one of the main culprits that contributes to excessive expert intervention and therefore, the need for greater automation level in interactive method has been echoed (Shim et al., 2009; Bae et al., 2009).

The improved random walks with automated seeds has demonstrated greater efficiency and reproducibility than manual method in segmenting knee cartilage. The improved semi-automated method has significantly decreased the processing time for normal cartilage segmentation by 47.5% for observer 1 and 44% for observer 2. Besides, the improved semi-automated method has also improved the efficiency of diseased cartilage segmentation by 48.1% for observer 1 and 30.3% for observer 2. The improvement in efficiency was attributed to the implementation of computer generated seeds that helped minimize the degree of expert intervention to cartilage seeds.

Besides, semi-automated method was more reproducible than manual method. Given that manual method involved laborious outlining of cartilage boundary, the method was sensitive to anatomical variation and human bias. Implementation of computer-aided algorithms, on the other hand, automated the boundary delineation process. Moreover, random walks algorithm has demonstrated excellent technical properties like robustness to image noise and weak boundary problem (Grady, 2006) compared to existing semi-automated algorithms like graph cuts (Boykov and Jolly, 2001), livewire (Gougoutas et al., 2004) and active contour (Kass et al., 1988). In addition, implementation of computer generated seeds did not compromise the high reproducibility promised by traditional semi-automated method; indicating the simplified labelling procedures has succeeded to preserve crucial expert guidance feature to the most needed part.

Nonetheless, our study has several limitations. First, we did not measure the volumetric changes of cartilage as reported in previous OA studies. In this works, we mainly focused on the development of an enhanced semi-automated segmentation method that simplified the labelling procedures. Second, the dataset size used in this study was relatively small and larger sample size should be considered in future works.
Conclusion

The improved semi-automated method represents a paradigm shift from existing segmentation techniques by incorporating extra automation to expert initialization. Therefore, the proposed method is able to increase the segmentation efficiency without compromising high reproducibility with manual segmentation’s results. In addition, random walks method has better technical properties compared to existing methods. As a result, the combination contributes to more user friendly segmentation environment.

Acknowledgement

The authors gratefully acknowledged the financial support provided by Universiti Kuala Lumpur, British Malaysian Institute (UniKL BMI).

Author's Contributions

Hong Seng Gan: Drafting of manuscript, Experimental planning and setup, Data acquisition, Data analysis.
Amir Khairil Sayuti: Drafting of manuscript, Data acquisition, Experimental planning and setup.

Conflict of Interest

All authors declare no conflict of interest.

References


