Segmenting the Femoral Head and Acetabulum in the Hip Joint Automatically Using a Multi-Step Scheme

Ji WANG†, Yuanzhi CHENG†(a), Yili FU†, Shengjun ZHOU†, Nonmembers, and Shinichi TAMURA††, Fellow

SUMMARY  We describe a multi-step approach for automatic segmentation of the femoral head and the acetabulum in the hip joint from three dimensional (3D) CT images. Our segmentation method consists of the following steps: 1) construction of the valley-emphasized image by subtracting valleys from the original images; 2) initial segmentation of the bone regions by using conventional techniques including the initial threshold and binary morphological operations from the valley-emphasized image; 3) further segmentation of the bone regions by using the iterative adaptive classification with the initial segmentation result; 4) detection of the rough bone boundaries based on the segmented bone regions; 5) 3D reconstruction of the bone surface using the rough bone boundaries obtained in step 4) by a network of triangles; 6) correction of all vertices of the 3D bone surface based on the normal direction of vertices; 7) adjustment of the bone surface based on the corrected vertices. We evaluated our approach on 35 CT patient data sets. Our experimental results show that our segmentation algorithm is more accurate and robust against noise than other conventional approaches for automatic segmentation of the femoral head and the acetabulum. Average root-mean-square (RMS) distance from manual reference segmentations created by experienced users was approximately 0.68 mm (in-plane resolution of the CT data). key words: hip joint, osteoarthritis, mathematical morphology, vertex normal, threshold selection

1. Introduction

Segmentation of the pelvis and the femur from three dimensional (3D) CT images is essential for computer assisted surgical planning, intraoperative navigation, and postoperative assessment [1]. The hip joint is a ball and socket joint, consisting of the femur and the pelvic bone. A normal hip joint is formed by the round ball-like upper end of the femur, the femoral head, in a socket-like cavity, the acetabulum, in the pelvic bone. The head of the femur is covered by a layer of smooth cartilage and is reinforced in its position by very powerful ligaments. CT scan of the hip joint usually reveals a space between the ball and the socket because of the very low density of cartilage on X-ray CT. This space is approximately 1/4-in wide and fairly even in outline in a normal hip. However, clinical complications such as osteoarthritis and congenital deformities, technical limitations of CT imaging like inherent blurring and partial volume effect, and the narrow interbone regions and the presence of blood vessels in the bone tissues often result in textured regions, as well as weak and diffused boundaries in pelvis and femur images.

Threshold-based techniques have been widely used in segmentation of the bone from CT images [2]–[4]. This type of segmentation is fairly successful in general since the CT values of bone are greater than those of the surrounding soft tissues. However, bone segmentation by a global threshold technique fails to work well in some regions like the hip joint where individual bone structures such as the acetabulum and the femoral head are physically adjacent. This leads to the well-known difficulty in segmentation of the two structures. Mathematical morphology is also a well-known technique used in bone segmentation. For example, images can be segmented into visually sensible regions by finding the watershed regions in a gradient magnitude image [5], [6]. Nevertheless, this approach has an over-segmentation problem, which has led to a number of studies on merging watershed regions to obtain objects of interest [7]. Recently, segmentation of the pelvic bone from 3D images using a statistical shape model has been reported [8], [9]. However, a statistical shape model requires a large number of training data sets.

Our aim is the development of an accurate and robust 3D segmentation scheme for the femoral head and the acetabulum in the hip joint from CT images. To our knowledge, the feasibility of all previously reported techniques for segmentation of the femoral head and the acetabulum has not been adequately investigated in actual patients. We also aim to verify the performance of the new technique in actual clinical cases to develop an integrated approach that can be employed in the clinical practice. The segmentation approach presented in this paper is specifically developed for CT images of the femoral head and the acetabulum but is applicable to other skeletal sites as well.

2. Overview of the Method Presented

Our segmentation approach consists of four modules: preprocessing, initial segmentation, iterative adaptive classification and correction.

The preprocessing module generates valley-emphasized images and is described in Sect. 3.2. The valley-emphasized images are constructed by subtracting valleys from the original images so that intensities between the femoral head and the acetabulum stand out. The second module is based on conventional techniques including the initial thresholding and binary morphological operations and is described in Sect. 3.3. In this module, initial segmen-

Copyright © 2012 The Institute of Electronics, Information and Communication Engineers
3. Materials and Methods

3.1 Data Sets

In the present study, we evaluated our approach on 35 patient data sets (70 hip joints; age range, 32-67 years; mean age, 49 years). A GE Toshiba CT machine was used with a field of view (FOV) of 320 mm (512 × 512 matrix, pixel size = 0.68 mm) and slice interval of 1.5 mm. The number of slices per data set ranged from 85 to 95 including the whole hip and pelvis bone. These patients were exhibited a wide range of bone pathology and morphometric variation, which increased the difficulty of automatic segmentation.

3.2 Preprocessing

The preprocessing module generates valley-emphasized images. The valley-emphasized images are constructed by subtracting valleys from the original images so that intensities between the femoral head and the acetabulum stand out.

Valley and gradients are two distinctive features in medical imaging analysis. However, valleys are more suitable features than gradients in separating two adjacent objects because they appear in the middle of two adjacent objects while gradients appear at the boundaries of each object. A valley is a dark blob, a region with significantly darker intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out.

Valley and gradients are two distinctive features in medical imaging analysis. However, valleys are more suitable features than gradients in separating two adjacent objects because they appear in the middle of two adjacent objects while gradients appear at the boundaries of each object. A valley is a dark blob, a region with significantly darker intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out.

Valley and gradients are two distinctive features in medical imaging analysis. However, valleys are more suitable features than gradients in separating two adjacent objects because they appear in the middle of two adjacent objects while gradients appear at the boundaries of each object. A valley is a dark blob, a region with significantly darker intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities relative to the surroundings [10]. As the narrow inter-bone region between the femoral head and the acetabulum stand out. The narrow inter-bone region pixel intensities rela...
higher peaks corresponded to image of intensities of cortical and spongy bones in the femoral head and the acetabulum. The lower peak corresponded to that of the surrounding soft tissues. Two Gaussian curves were fitted to the highest intensity values of these two separated peaks. Using these Gaussian curves, we calculated a threshold value based on bimodal distribution, and binarized the smoothed images [12]. As shown in Fig. 2, the intersection of the two curves was selected as a threshold value. A binary image was constructed by the selected threshold value. Each pixel of the valley-emphasized CT image was divided into the bone and non-bone classes.

(2) Binary Morphological Operations. Binarizing the data set by the above method might create holes in the resultant binary data. In addition, some unwanted structures may appear as small regions in the data sets. We employed 3D binary morphological operations to close open regions and to fill holes of the binary data obtained after threshold. In the following, we will discuss in short the binary morphological operators. We limit the discussion to three dimensions. Furthermore only sets and functions defined in the Euclidean grid \( \mathbb{Z}^3 \) will be considered.

For each voxel \( v = (z, y, x) \) of the binary 3D image, mathematical morphology describes an object \( Y \) in a binary 3D image as a set in a 3D grid:

\[
Y = \{v : f(v) = 1, v = (z, y, x) \in \mathbb{Z}^3\} \tag{6}
\]

where \( f \) is called the characteristic function of \( Y \). On the other hand, the object background \( Y^b \) can be defined as follows:

\[
Y^b = \{v : f(v) = 0, v = (z, y, x) \in \mathbb{Z}^3\} \tag{7}
\]

Morphological operations are essential interactions between a user-defined object \( G \) called structuring element and the objects that are present in the image. The structuring elements are in most cases simple geometric shapes: spheres, cubes, diamonds and so on. In this study, the given structuring elements are defined as \( 3 \times 3 \times 3 \) diamond shape [13]. Figure 3 shows a 3D structuring element containing 6 neighborhoods. The symmetric of \( G \) with respect to the origin \((0, 0, 0)\) is denoted by \( G^s\).

\[
G^s = \{-v : v \in G\} \tag{8}
\]

whereas the translate of of \( G \) by a vector \( d \) is denoted by \( G_d \)

\[
G_d = \{v + d : v \in G\} \tag{9}
\]

The two elementary morphological operations are dilation

\[
Y \oplus G^s = \{v \in \mathbb{Z}^3 : G_e \cap Y \neq \emptyset\} \tag{10}
\]

and erosion

\[
Y \ominus G^s = \{v \in \mathbb{Z}^3 : G_e \subset Y\} \tag{11}
\]

Application of dilation on an eroded binary 3D image does not lead to the initial image; that is, the two operators are nonreversible. Successive application of erosion and dilation is termed opening:

\[
Y_G = (Y \ominus G^s) \oplus G \tag{12}
\]

Opening smoothes the surface of an object, smearing sharp spurs on the object boundary. It also cuts small isthmuses, thus separating the objects united by such isthmus. The operation resulting from the inverse succession of procedures, namely dilation followed by erosion, is called closing:

\[
Y_G^c = (Y \oplus G^s) \ominus G \tag{13}
\]

Closing fills small holes and gulfs on the 3D object surface. It also unites objects that are close to each other.

The steps of the initial segmentation are shown in Fig. 4. After performing the initial segmentation, some unwanted structures (ellipse area and rectangular area in Figs. 4(b), (d) and (e)) and holes (triangular area in Fig. 4(d)) maybe appear in the resultant binary data. This is due to weak and diffused boundaries in the acetabulum and the femoral head, technical limitations of CT imaging, and partial volume effect. By comparing Figs. 4(b) and (d), it can be seen that the binary images detected from the valley-emphasized images have a better results relative to those detected from the original images. It can save a lot of time in the following iterative adaptive process of 3D correlation to update voxels classification (See the ellipse regions in Figs. 4(b) and (d)). The segmentation result is used as the initial input to an iterative adaptive threshold scheme that is described in the next section.
1) Compute \( E_B \) from the current segmentation (\( B \) and \( T \)).
2) For each voxel \( v \) in \( E_B \):
   - assume that the voxels in \( E_B \) come from a mixture of two Gaussian distributions (bone and tissue) having means and standard deviations \((\mu_B, \sigma_B)\) and \((\mu_T, \sigma_T)\), respectively.
   - classify \( W(v) \) using the Bayes decision rule [15], and add \( v \) to the error class \( R \) if \( v \) is classified as tissue.
3) Update the current segmentation: \( B \leftarrow B/R, T \leftarrow T/R \).
4) Iterate until convergence.

After performing the iterative adaptive threshold algorithm with the initial segmentation, each of the separated regions is a closed region on each slice of CT images. All components that are smaller than a specified threshold value (8 voxels) are discarded as noise, and the remaining components where identified as bone regions, as illustrated in Fig. 5 (h). The outer contours of the bone can be identified by judging 4-neighborhood structure of each pixel in the binary images slice by slice, which is similar to compute boundary pixels from its directly connected 3D neighbors. This operation is executed on each 2D slice separately, rather than in 3D. The reason that we restrict the one to 2D is that single pixel contours of the bone can be derived fast and accurately while the mistaken outer contours of the bone detected in 3D are avoided, the segmentation procedure is shown in Fig. 5.

### 3.5 Correction Method based on the Normal Direction of Vertices

The segmentation schemes described in the previous section do not yield a satisfactory segmentation yet. Therefore, we finally applied an automatic correction technique to refine the preliminary contour of the acetabulum and the femur. The refinement consists of three major steps: 1) a triangulation of the bone surface obtained in Sect. 3.4; 2) correction of vertices based on the vertex normal of the triangulated surface; and 3) an adjustment of the triangulated surface.

The zero-crossing of the second derivative (ZSD) method is a widely used technique in the edge detection processing which is accurate and precise. The first derivative of the image function should have an extremum at the position corresponding to the edge in the image, and so the second derivative should be zero at the same position. However, in order to compute the second derivative robustly, one possibility is to smooth an image first to reduce noise which increases the computation time inevitably.

Our approach based on the normal direction of vertices is a modified ZSD method, which computes the maximum value of the first derivative in the neighborhood of the boundary voxels instead of the zero-crossings of the second derivative along the normal direction. In practice, our method requires lower computational efforts and is more robust against noise. Our 3D correction method is divided into...
three major steps.

1) A triangulation of the bone surface. The 3D bone surface is reconstructed by using the rough bone boundaries obtained in Sects. 3.2 - 3.4. The 2D contours are processed to form a 3D surface representation (3D reconstruction) of the bone, based on a network of triangles, by connecting the points of one 2D contour with its closest counterparts on the consecutive 2D contours.

2) Correction of vertices based on the vertex normal of the triangulated surface. There are some obvious advantages to use the vertex normal direction in image processing. First, the vertex normal near the boundary is more stable than the intensity of voxel near the boundary; second, the vertex normal direction encodes 3D geometric information, which is very useful in image processing. Let us consider $N$ the number of vertices of the triangulated surface. For a vertex, it is the number of vertices of the triangulated surface. For a vertex $S_i(i = 1, 2, \ldots, N)$, its normal is given by [16]

\[
n_j = \frac{\sum_{j \in F(i)} \theta_p n_i(F_j)}{\sum_{j \in F(i)} \theta_p n_i(F_j)} \pm (\cos \alpha, \cos \beta, \cos \gamma) \quad (14)
\]

where $F$ is the triangulated facets, $j$ is its index; $F(i)$ is face index set containing the vertex $S_i$; $n(S_i)$ is the normal of incident face $F_j$, and $\theta_j$ is the angle formed by the edges $F_j$ incident to $S_i$. The parameters $\alpha$, $\beta$, and $\gamma$ are the angles between the normal vector of each vertex $S_i$ and the $x$, $y$, and $z$-axes, respectively. Although there are various ways of estimating the surface normal, we choose this one because it is fast, local, and achieves good results.

3) An adjustment of the triangulated surface. Let $V_i$ be the $i$th vertex in the triangulated surface. We select the four nearest voxels along the vertex $V_i$ normal $n_i$ centered at the vertex $V_i$, which are expressed as $V_{i-2}$, $V_{i-1}$, $V_{i+1}$, $V_{i+2}$ (See Fig. 6). Meanwhile, $V_i$ is denoted as $V_{i,0}$. Thus, a total of five voxels, including the vertex $V_{i,0}$, are $V_{i-2}$, $V_{i-1}$, $V_{i,0}$, $V_{i+1}$, and $V_{i,2}$, respectively. $C_{i-2}$, $C_{i-1}$, $C_{i,0}$, $C_{i,1}$, and $C_{i,2}$ are the central positions of five voxels in the coordinate system for $V_{i-2}$, $V_{i-1}$, $V_{i,0}$, $V_{i+1}$, and $V_{i,2}$, respectively. The first order directional derivatives of five voxels along the vertex $(V_{i,0})$ normal $n_i$ are expressed as $f'(V_{i,k})_n$, where $k = -2, -1, 0, 1, 2$. The first and second maximum values computing $\|f'(V_{i,k})_n\|$ are given by

\[
V_{i,p} = \max V_{i,k} \|f'(V_{i,k})_n\| \quad (15)
\]

\[
V_{i,q} = \max V_{i,k} \|f''(V_{i,k})_n\| \quad (16)
\]

An adjustment of the triangulated surface is summarized in Algorithm 1. Figure 7 demonstrates two examples of Algorithm 1.

![Fig 6](image.png) Selection of four voxels along the normal direction of vertex $V_{i,0}$.

![Fig 7](image.png) Two examples of Algorithm 1.
Algorithm 1: Adjustment of the triangulated surface

```
input : The triangulated surface.
output: The adjusted triangulated surface.

while iterations does not reach a certain number or there are obvious changes in each vertex. do
  for each Vi or Vq do
    Compute ni, V1i, and Vq;
    if V0 ≠ Vp and V0 ≠ Vq then
      if Ci0 = (Ci,p + Ci,q) / 2 then
        if ||f'(Vi,p)ni|| ≥ ||f'(Vi,q)ni|| then
          Replace V0 by Vi,p;
        else
          Replace V0 by Vi,q;
        end
      else
        if ||Ci,q − Ci,0|| ≤ ||Ci,p − Ci,0|| then
          Replace V0 by Vi,q;
        else
          Replace V0 by Vi,p;
        end
      end
    end
  end
end
```

Although there are two rules which can be chosen in our algorithms: one is the vertex is replaced with the position having the maximum directional derivative along the vertex normal direction, the other is that replacing the vertex with the nearest position which has a relative larger directional derivative. Considering the smooth of the adjusted triangulated surface, partial volume effect, the effect of cortical bone thickness and the narrow inter-bone regions, the latter rule is applied in our study.

4. Experiment Results

In the following section, we evaluated our segmentation algorithm by comparing with some other segmentation methods.

4.1 Comparison with Global-Threshold Segmentation

Since the intensity values of bone, especially for cortical bone, are statistically higher than the intensity of the surrounding tissue, an intensity threshold can be detected such that the intensities of most pixels of bone are greater than the threshold value, so global threshold is a commonly used method in the automated segmentation of bone from CT images [4]. However, due to the narrow inter-bone regions as well as weak and diffused boundaries in pelvis and femur, it is seldom possible to find a threshold that is less than values of all bone and greater than values of all other tissues. Figure 4(b) shows the segmentation results on three slices of images when the threshold was set to be 1150 Hounsfield units, some unclosed contours and false connections appeared during this process.

Global threshold method was also used in our approach to preprocess the volumetric CT images into the bone and non-bone classes. The differences is that we performed this operation on valley-emphasizes instead of on original images. Figure 4(d) shows our threshold segmentation results, which were better than the results of Fig. 4(b). Figure 8 shows the 3D visualization results with global threshold segmentation after discarding noisy voxels. (a) The posterior view. (b) The anterior view. Black arrows point out the false connections.

4.2 Comparison with Zoroofi’s Method

Zoroofi et al. [17] described a multi-step method for automatic segmentation of the pelvis and the femur from volumetric CT images. Their segmentation approach consisted of four major steps: 1) detecting bone tissues from CT images by conventional techniques including histogram-based threshold and binary morphological operations; 2) estimating initial boundary of the femoral head and the joint space between the acetabulum and the femoral head by a new approach utilizing the constraints of the greater trochanter and the shapes of the femoral head; 3) enhancing the joint space by a Hessian filter; and 4) refining the rough boundary obtained in step 2) by a moving disk technique and the filtered images obtained in step 3).

We tested and quantified the performance of Zoroofi’s method on numerous slices of the acetabulum and the femoral head images, and found that it did a good job on most slices of CT images even for some difficult slices. However, there are some cases that Zoroofi’s method results in poor segmentations and examples are displayed in Fig. 9. This is because the Hessian filter cannot perfectly enhance the joint space between the acetabulum and the femoral head.

4.3 Comparison with Manual Segmentation

Manually traced surfaces were used for the gold standard,
comparisons between the estimated and gold standard surfaces were performed. To evaluate the segmentation accuracy, two types of measurements were used: one was Jaccard similarity measure [18] and the other was root-mean-square (RMS) distance. The former was defined as , where $A$ and $B$ were the estimated region using our approach and the corresponding gold standard region, respectively. The latter is the RMS symmetric distance between the estimated and gold standard surfaces (boundaries). The RMS is given by the following equation:

$$RMS = \sqrt{\left(\frac{1}{N} \sum_{i=1}^{N} d_i^2\right)/N}$$

where $1 \leq i \leq N$ and $N$ is the number of surface points, and $d_i$ is the Euclidean distance between the point $i$ on the estimated surface and the corresponding point $i$ on the gold standard surface.

Figure 10 shows the representative results of the proposed segmentation technique and the manual segmentation for two typical cases. A comparison between the segmented region using our approach and the corresponding gold standard region was performed (Figs. 10 (e) and (f)). The average RMS distance between estimated and gold standard boundaries was 0.71 mm for Fig. 10(e) and 0.65 mm for Fig. 10(f).

We performed the accuracy comparison among the three segmentation approaches (global thresholding method, Zoroofi’s method and our method). Our method can automatically separate all regions of acetabulum and femoral head on 35 patient data sets. However, Global thresholding and Zoroofi’s methods cannot satisfactorily handle all the data sets. We manually handled 32% of the total slices for the global thresholding method and 8% of the total slices for the Zoroofi’s method.

Figure 11 shows accuracy evaluation results using 35 patient data sets for each of three different segmentation methods. Two types of measurements were used for accuracy evaluation. The results of Jaccard similarity measure are shown in Fig. 11 (a). The average Jaccard similarity measure ($N = 35$) was 78.3% (range, 72.6% – 82.7%) by global thresholding method, 92.7% (range, 88.7 – 98.2%) by Zoroofi’s method, and 97.1% (range, 95.9% – 98.3%) by our proposed method. The results of the RMS distance are shown in Fig. 11 (b). The average RMS distance error ($N = 35$) was 1.03 mm (range, 0.87 – 1.32 mm) by global thresholding method, 0.76 mm (range, 0.51 – 1.12 mm) by Zoroofi’s method, and 0.68 mm (0.42 – 1.03 mm) by our proposed method. It can be shown in Fig. 11 that the proposed segmentation method was the best among the three segmentation approaches.

Our method satisfactorily carried out the segmentation of the acetabulum and the femoral head on 35 data sets. In clinical cases with severe OA, however, our automatic technique cannot successfully classify all regions of acetabulum and femoral head. Therefore, manual processing is required at the moment. Figure 12 shows the view of a case with severe OA. Figure 12 (a) displays an illustrative case using our automatic method. In 3 slices of the hip joint, the acetabulum and the femoral head cannot be separated correctly. Black arrows point out the false connections. (b) Segmentation of a hip showing the separation of the femurs and the acetabulum after manually segmenting these 3 slices.
rectly. After manually segmenting these 3 slices, we obtain a good segmentation result (Fig. 12(b)).

5. Discussion

We have developed a multi-step segmentation method for the pelvis and femur, especially for the acetabulum and the femoral head from CT images. We evaluated our approach on 35 CT data sets. The segmentation accuracy was assessed in two ways. First, we compared the segmentation result with other comment segmentation methods, such as global threshold method [5] and Zoroofi’s method [17]. Second, we compared the segmentation result which was manually segmented and reviewed by one experienced body radiologist. We show that the proposed method leads to very accurate segmentations of the hip joint region.

All the methods have been programmed using Matlab 6.5 in an Intel (R) Core (TM) 2 Duo E6550 computer, 2.33 GHz, 2 G RAM. The processing time will be a little less in Microsoft Visual C++ compiler environment than that in MATLAB environment. However, for an experienced radiologist it is required to take more than two hours to segment the acetabulum and the femoral head manually in a scan with slice by slice, so our approach is fast and suitable for medical application.

In our experiments, the execution time for performing automatic segmentation on a personal computer was less than 17 min in MATLAB environment, which was mainly exhausted on iterative adaptive classification and normal direction correction operation. The processing time for the 35 patient datasets was about 10 s per slice. This execution time can be further reduced by optimizing the iterative convergence process.

In the accuracy analysis, the manual segmentation results were considered as ground truth. The segmentation result using our method was compared to the ground truth. The voxel dimensions of the CT data in our study were non-cubic: 0.68 × 0.68 × 1.5 mm$^3$. In order to obtain the accurate surface distance of the acetabulum and the femoral head from the ground truth segmentations, we resampled both images (automatic segmentation image using our method and manual segmentation image) at 1/2 number of in-plane resolution (0.68 mm) and 1/4 number of slice thickness (1.5 mm). The modified Sinc interpolation technique described in [19] was employed in our study to remove the unwanted Gibb’s ringing. The sampling interval in the interpolated CT data was 0.34 × 0.34 × 0.37 mm$^3$ in the x-, y-, and z- axes.

The status of the joints range from healthy to osteoarthritic according to the Kellgren-Lawrence index (KL$_i$) [20], a radiographic score established by X-rays between 0 - 4 where KL$_0$ = 0 is healthy, KL$_1$ = 1 is doubtful OA, KL$_2$ = 2 is minimal OA, KL$_3$ = 3 is moderate OA, and KL$_4$ = 4 is severe OA. In the present study, 61 hips have KL$_i$ ≤ 2 and 9 hips have KL$_i$ = 3. In clinical cases with KL$_i$ = 4, our 3D adaptive iterative classification approach cannot classify all regions of acetabulum and femoral head correctly, and manual processing is required at the moment (Fig. 12).

6. Conclusion

The proposed method based on a multi-step scheme can be used to accurately segment the femoral head and the acetabulum in the hip joint with an error on the order of an in-plane resolution. An extensive evaluation of the proposed method on 35 CT datasets revealed a high segmentation quality compared to other conventional approaches.

Automatic segmentation of the acetabulum and the femoral head in clinical cases with severe OA is a challenging task. Our future work will focus mainly on the development of post-processing algorithms for an improved segmentation of the acetabulum and the femur in the hip joint with severe OA.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China under Grant No.61073124; the National Basic Research Program of China (973) under Grant No.2011CB707701; the Project under Grant HIT-NSRIF, 2009149 and SKLRS-2009-ZD-03.

References

[1] N. Sugano, “Computer-assisted orthopedic surgery,” J. Orthop. Sci., vol.8, no.3, pp.442–448, 2003.
[2] T. Klinder, J. Ostermann, M. Ehm, A. Franz, R. Kneser, and C. Lorenz, “Automated model-based vertebra detection, identification, and segmentation in CT images,” Med. Image Anal., vol.13, no.3, pp.471–482, 2009.
[3] T.N. Hangartner, “Thresholding technique for accurate analysis of density and geometry in QCT, pQCT and microCT images,” J. Musculo skeletal and Neuronal Interactions, vol.7, no.1, pp.9–16, 2007.
[4] J. Dehmeshki, X. Ye, H. Amin, M. Abaie, X. Lin, and S.D. Qanadli, “Volumetric quantification of atherosclerotic plaque in CT considering partial volume effect,” IEEE Trans. Med. Imaging, vol.26, no.3, pp.273–282, 2007.
[5] A. Bieniek and A. Moga, “An efficient watershed algorithm based on connected components,” Pattern Recognit., vol.33, no.6, pp.907–916, 2000.
[6] H. Sun, J. Yang, and M. Ren, “A fast watershed algorithm based on chain code and its application in image segmentation,” Pattern Recognit. Lett., vol.26, no.5, pp.1266–1274, 2005.
[7] P.S.U. Adiga and B.B. Chaudhuri, “An efficient method based on watershed and rule-based merging for segmentation of 3-D histopathological images,” Pattern Recognit., vol.34, no.7, pp.1449–1458, 2001.
[8] T. Okada, K. Yokota, M. Hori, M. Nakamoto, H. Nakamura, and Y. Sato, “Construction of hierarchical multi-organ statistical atlases and their application to multi-organ segmentation from CT images,” Proc. MICCAI 2008, New York, USA, Part I. LNCS, vol.5241, pp.502–509, Sept. 2008.
[9] B. Hass, T. Coradi, M. Scholz, P. Kunz, M. Huber, U. Oppitz, L. Andre, V. Lengkeek, D. Huyskens, A. van Esch, and R. Reddick, “Automatic segmentation of thoracic and pelvic CT images for radiotherapy planning using implicit anatomic knowledge and organ-specific segmentation strategies,” Phys. Med. Biol, vol.53, pp.1751–1771, 2008.
[10] P. Maragos, “Morphological filtering for image enhancement and feature detection,” in Handbook of image and video processing, ed. A.C. Bovik, 2nd ed., pp.135–136, Academic Press, 2005.

[11] N. Otsu, “A threshold selection method from gray level histograms,” IEEE Trans. Syst. Man. Cybern. A, Syst. Humans, vol.9, no.1, pp.62–66, 1979.

[12] R. Jain, R. Karsuri, and B.G. Schunck, Machine Vision, McGraw-Hill, New York, 1995.

[13] D.F. Rogers, Procedural Elements for Computer Graphics, McGraw-Hill, New York, 1985.

[14] C.H. Yan, Measuring changes in local volumetric bone density: new approaches to quantitative computed tomography, Ph.D. Thesis, Stanford University, 1998.

[15] J. Berger, Statistical Decision Theory and Bayesian Analysis, Springer, Berlin, 1985.

[16] S. Jin, B. Lewis, and D. West, “A comparison of algorithm for vertex normal computation,” Vis. Comput., vol.21, no.1, pp.71–82, 2005.

[17] R.A. Zoroofi, Y. Sato, T. Sasama, T. Nishii, N. Sugano, K. Yonenobu, H. Yoshikawa, T. Ochi, and S. Tamura, “Automated segmentation of acetabulum and femoral head from 3-D CT images,” IEEE Trans. Inf. Technol. B., vol.7, no.4, pp.329–343, 2004.

[18] P. Jaccard, “The distribution of the flora in the alpine zone,” New Phytologist, vol.11, no.2, pp.37–50, 1912.

[19] Y. Sato, C.F. Westin, A. Bhalariao, S. Nakajima, N. Shiraga, S. Tamura, and R. Kikinis, “Tissue classification based on 3D local intensity structures for volume rendering,” IEEE Trans. Vis. Comput. Graph., vol.6, no.2, pp.160–180, April-June, 2000.

[20] J. Kellgren and J. Lawrence, “Radiological assessment of osteoarthrosis,” Ann. Rheumatic Diseases, vol.16, no.4, pp.494–502, 1957.

Ji Wang is currently studying in School of Mechatronics Engineering, Harbin Institute of Technology, China.

Yuanzhi Cheng received the Ph.D. degree from Osaka University Medical School, Osaka, Japan in 2007. He is currently an Associate Professor in School of Mechatronics Engineering, Harbin Institute of Technology. His research interests are concentrated on medical image processing, computer-integrated surgical systems and surgical informatics.

Yili Fu received his Ph.D. in Mechatronics from Harbin Institute of Technology (HIT), China in 1996. He is currently a Professor of Robotics Institute at HIT. He serves as the Associate Editor of International Journal of Humanoid Robots since 2003. He has published over 150 papers in journals and conference proceedings. His research interests include biomechanical engineering, medical robotics and medical imaging.

Shengjun Zhou is currently completing the Ph.D. in School of Mechatronics Engineering, Harbin Institute of Technology, China.

Shinichi Tamura received the B.S., M.S., and Ph.D. degrees in electrical engineering from Osaka University in 1966, 1968, and 1971, respectively. He is currently a Professor in the Center for Advanced Medical Engineering and Informatics, Osaka University. Dr. Tamura is a Fellow of IEEE.