Fully Automatic Segmentation of Hip CT Images via Random Forest Regression-based Atlas Selection and Optimal Graph Search-based Surface Detection

Chengwen Chu[1], Junjie Bai[2], Li Liu[1], Xiaodong Wu[2], and Guoyan Zheng[1]

[1] Institute for Surgical Technology and Biomechanics, University of Bern, Switzerland {chengwen.chu, li.liu}@istb.unibe.ch, guoyan.zheng@ieee.org
[2] Dept. Electrical and Computer Engineering, The University of Iowa, USA {junjie-bai, xiaodong-wu}@uiowa.edu

Abstract. Automatic extraction of surface models of both pelvis and proximal femur of a hip joint from 3D CT images is an important and challenging task for computer assisted diagnosis and planning of periacetabular osteotomy (PAO). Due to the narrowness of hip joint space, the adjacent surfaces of the acetabulum and the femoral head are hardly distinguishable from each other in the target CT images. This paper presents a fully automatic method for segmenting hip CT images using random forest (RF) regression-based atlas selection and optimal graph search-based surface detection. The two fundamental contributions of our method are: 1) An efficient RF regression framework is developed for a fast and accurate landmark detection from the hip CT images. The detected landmarks allow for not only a robust and accurate initialization of the atlases within the target image space but also an effective selection of a subset of atlases for a fast atlas-based segmentation; and 2) 3-D graph theory-based optimal surface detection is used to refine the extraction of the surfaces of the acetabulum and the femoral head with the ultimate goal to preserve hip joint structure and to avoid penetration between the two extracted surfaces. Validation on 30 hip CT images shows that our method achieves high performance in segmenting pelvis, left proximal femur, and right proximal femur with an average accuracy of 0.56 mm, 0.61 mm, and 0.57 mm, respectively.

1 Introduction

Developmental dysplasia of hip (DDH) is a congenital defect that seriously affects young people nowadays. In many treatment procedures for patients with DDH, periacetabular osteotomy (PAO) recently becomes a common surgical intervention [1], aiming to improve ability of weight bearing and stability of the diseased hip joint. To reach this goal, knowing acetabular coverage, which is defined as a ratio between the femoral head surface covered by the acetabulum and the complete femoral head surface, is important for operative planning for

PAO. For this purpose, we need to extract surface models of both the pelvis and the proximal femur from hip CT images.

Automatic extraction of the surface models of both the pelvis and the proximal femur from hip CT images comprises two key steps. Firstly, both anatomical structures have to be detected in the target volume data and secondly, both models need to be segmented. Furthermore, the fact that the two structures compose a hip joint should not be neglected. Otherwise, the resultant models may penetrate each other due to the narrowness of the hip joint and hence do not represent a true hip joint.

For detection, reported methods in literature address the problem either by assuming an user-supplied initialization [2, 3] or by using Generalized Hough Transform (GHT) [4, 5]. For segmentation, both multi-atlas-based segmentation methods [6–8] and statistical shape model (SSM)-based segmentation methods [2–5, 9, 10] are proposed. Here we define an atlas as a pair of data consisting of a CT volume and its corresponding segmentation. Given a set of atlases, atlas-based segmentation methods segment a target volume by registering the atlases to the volume first, followed by a label fusion process. Multi-atlas-based segmentation methods may be applicable for extraction of surface models of individual structures of the hip joint, but they cannot guarantee the preservation of the hip joint space and the prevention of the penetration of the extracted surface models. The other segmentation option is the SSM-based methods, which perform an adaption of the SSM to the target image data. Similar to atlas-based methods, conventional SSM-based methods are difficult, if not impossible, to guarantee the preservation of the hip joint structure [2–4, 9]. This problem is recently addressed by introducing an articulated statistical shape model (aSSM) [5]. Another solution is to simultaneously detecting both surfaces of the adjacent structures based on graph optimization theory [10]. By incorporating prior knowledge about spatial relationship in the graph optimization, the adjacent surfaces can be segmented without penetration to each other.

In this paper, we propose a two-stage automatic hip CT segmentation method. In the first stage, we use a multi-atlas-based method to segment the regions of the pelvis and the bilateral proximal femurs. An efficient random forest (RF) regression-based landmark detection method is developed to detect landmarks from the target CT images. The detected landmarks allow for not only a robust and accurate initialization of the atlases within the target image space but also an effective selection of a subset of atlases for a fast atlas-based segmentation. In the second stage, we refine the segmentation of the hip joint area using graph optimization theory-based multi-surface detection [11, 12], which guarantees the preservation of the hip joint space and the prevention of the penetration of the extracted surface models with a carefully constructed graph. Different from the method introduced in [10], where the optimal surfaces are detected in the original CT image space, here we propose to first unfold the hip joint area obtained from the multi-atlas-based segmentation stage using a spherical coordinate transform and then detect the surfaces of the acetabulum and the femoral head in the unfolded space. By unfolding the hip joint area using the spherical coordinate

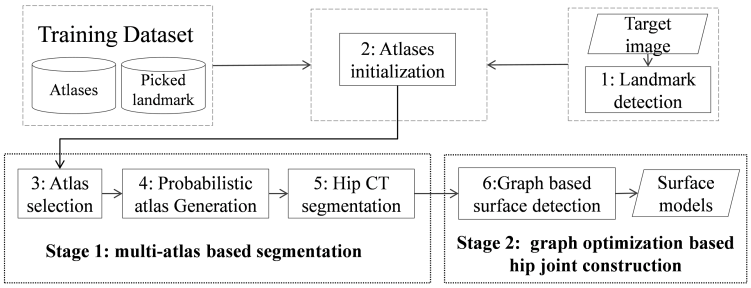

Fig. 1. The flowchart of our proposed segmentation method

transform, we convert the problem of detection of two half-spherically shaped surfaces of the acetabulum and the femoral head in the original image space to a problem of detection of two terrain-like surfaces in the unfolded space, which can be efficiently solved using the methods presented in [11, 12]. Fig. 1 presents a schematic overview of the complete workflow of our method.

2 Multi-atlas based hip CT segmentation

2.1 Landmark detection by fast random forest regression

Basic algorithm. We have a separate RF landmark detector for each landmark. During training, in each training image, we sample a set of image volumes around the ground-truth landmark position which is known. Each sampled volume is represented by its visual feature $\mathbf{f}_i \in \mathbb{R}^{d_f}$ and the displacement $\mathbf{d}_i \in \mathbb{R}^3$ from its center to the landmark (Fig. 2(a)). Let us denote all the sampled volumes in all training images as $\{P_i = (\mathbf{f}_i, \mathbf{d}_i)\}_{i=1...N}$ (Fig. 2(b)). The goal is then to learn a mapping function $\phi : \mathbb{R}^{d_f} \rightarrow \mathbb{R}^3$ from the feature space to the displacement space. Principally, any regression method can be used. In this paper, similar to [13, 14], we utilize the random forest regressors [15].

Once the regressor is trained, given a new image (Fig. 2(c)), we randomly sample another set of volumes $\{P'_k = (\mathbf{f'}_k, \mathbf{c'}_k)\}_{k=1...N'}$ all over the image (or a region of interest if an initial guess of the landmark position is known), where $\mathbf{f'}_k$ and $\mathbf{c'}_k$ are the visual feature and center coordinate of the kth volume, respectively (Fig. 2(d)). Through the trained regressor ϕ, we can calculate the predicted displacement $\mathbf{d'}_k = \phi(\mathbf{f'}_k)$, and then $\mathbf{d'}_k + \mathbf{c'}_k$ becomes the prediction of the landmark position by a single volume P'_k (Fig. 2(e)). Note that each tree in the random forest will return a prediction. Therefore, supposing that there are t trees in the forest, we will get $N' \times t$ predictions. These individual predictions are very noisy, but when combined, they approach an accurate prediction. To this end, we consider each single vote as a small Gaussian distribution. We developed a fast probability aggregating algorithm as described below to add these distributions to get a soft probability map called *response volume* which

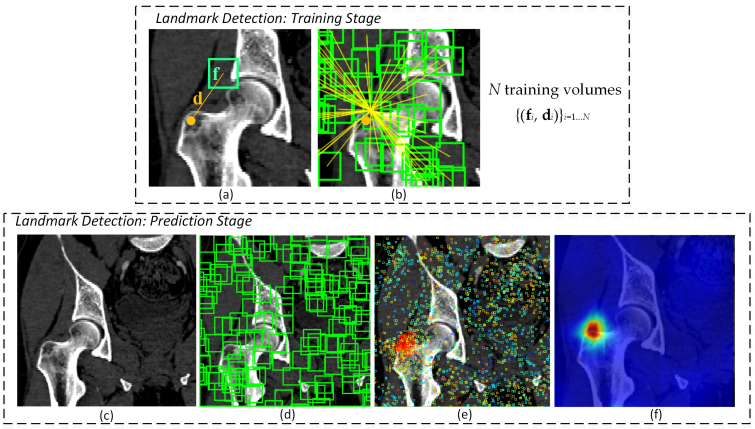

Fig. 2. The RF training and landmark detection. Illustration on coronal slice for easy understanding. (a) A volume sampled around the true landmark position. (b) Multiple sampled training volumes from one atlas. (c) A target image. (d) Multiple sampled test volumes over target image. (e) Each volume gives a single vote for landmark position. (f) Response volume calculated using improved fast Gaussian transform.

gives, for every position of the CT volume, its probability of being the landmark (Fig. 2(f)).

Fast probability aggregation. As described above, $N' \times t$ predictions are produced to detect each landmark. We consider each prediction a Gaussian model $\mathcal{N} \sim (\bar{\mathbf{d}}_k, \Sigma(\mathbf{d}_k))$, where $\bar{\mathbf{d}}_k$ and $\Sigma(\mathbf{d}_k) = diag(\sigma_{k,x}^2\ \sigma_{k,y}^2\ \sigma_{k,z}^2)$ are mean and covariance (which can be calculated from the displacements of the training samples that arrived at particular leaf node). All the $N' \times t$ predictions are accumulated to compute the likelihood of being a true landmark position for all M voxeles in the image. This finally yields a response volume for each landmark. Once the response volume has been obtained for each landmark, the position mode is selected as the landmark position.

The computational time of landmark prediction is mainly on multivariate Gaussian accumulation which is usually computed using

$$G(\mathbf{y}_i) = \sum_{k}^{N' \times t} \frac{1}{\sqrt{(2\pi)^3 |\Sigma(\mathbf{d}_k)|}} exp(-\frac{1}{2}(\mathbf{d}_{\mathbf{y}_i} - \bar{\mathbf{d}}_k)^T \Sigma(\mathbf{d}_k)^{-1}(\mathbf{d}_{\mathbf{y}_i} - \bar{\mathbf{d}}_k)) \quad (1)$$

where $\mathbf{d}_{\mathbf{y}_i} = \mathbf{y}_i - \mathbf{x}_k$, $\mathbf{y}_i$ is a voxel in target image and $\mathbf{x}_k$ is the center of volume k. For all of the N_l landmarks, such calculation will result in prohibitively expensive computation time of $O(M \times N' \times t \times N_l)$ on a 3D CT image with M voxels. In

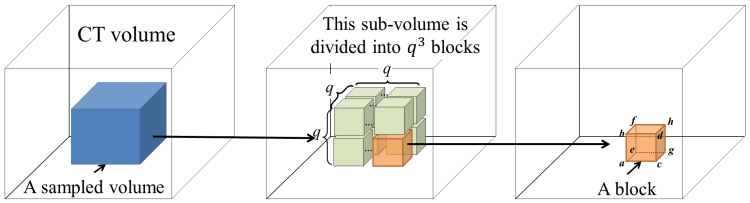

Fig. 3. A schematic view illustrating how to compute the visual feature of a sampled sub-volume for RF training and regression. Left: a sub-volume is sampled from a hip joint CT volume. Middle: we subdivide the sampled sub-volume into $q \times q \times q$ blocks. Right: for each block, we compute its mean and variance using the integral image technique

this paper, we propose to approximate Eq.1 by:

$$G(\mathbf{y}_i) = \sum_{k}^{N' \times t} W_k \cdot e^{(\|\mathbf{d}_{\mathbf{y}_i} - \bar{\mathbf{d}}_k\|^2 / h^2)} \tag{2}$$

Here we rewrite the Eq. 1 by introducing a constant kernel size of h, and moving the constrains of the variance out of the exponential part by introducing a weight $W_k = 1/\sigma_{k,x}\sigma_{k,y}\sigma_{k,z}$. With such an approximation, we develop an efficient probability aggregation strategy based on the Improved Fast Gaussian Transform (IFGT) [16] to calculate the response volumes with highly reduced time of $O((M + N' \times t) \times N_l)$.

Visual feature. As for the visual feature over the sampled sub-volume, we use mean and variance of intensities in a small volume obtained by subdividing the sampled sub-volume. In this paper, we subdivide each sampled sub-volume into a grid of $q \times q \times q$ blocks (see Fig. 3 for details). To accelerate the feature extraction within each block, we use the well-known integral image technique as introduced in [17]. Details about how to compute the integral image of a quantity can be found in [17]. The quantity can be the voxel intensity value or any arithmetic computation on the intensity value. Advantage of using integral image lies in the fact that once we obtain an integral image of the quantity over the complete hip CT volume, the sum of the quantity in any sub-volume can be calculated quickly in constant time O(1) regardless of the size of the volume [17]. Here we assume that we already computed the integral image of the voxel intensity I and the integral image of the squared voxel intensity S of the complete hip CT volume using the technique introduced in [17]. We then compute the mean $E[X]$ and the variance $Var(X)$ of the intensity value of any block (Fig. 3, right) as:

$$\begin{cases} E\left[X\right] = \left(I\left(\mathbf{h}\right) - I\left(\mathbf{d}\right) - I\left(\mathbf{f}\right) - I\left(\mathbf{g}\right) + I\left(\mathbf{b}\right) + I\left(\mathbf{c}\right) + I\left(\mathbf{e}\right) - I\left(\mathbf{a}\right)\right)/N \\ E\left[X^2\right] = \left(S\left(\mathbf{h}\right) - S\left(\mathbf{d}\right) - S\left(\mathbf{f}\right) - S\left(\mathbf{g}\right) + S\left(\mathbf{b}\right) + S\left(\mathbf{c}\right) + S\left(\mathbf{e}\right) - S\left(\mathbf{a}\right)\right)/N \\ Var\left(X\right) = E\left[X^2\right] - \left(E[X]\right)^2 \end{cases}$$

where $\{\mathbf{a}, \dots \mathbf{h}\} \in \mathcal{R}^3$ are the eight vertices of a block and N is the number of voxels within the block, as shown in Fig.3, right.

2.2 Atlas initialization and atlas-based segmentation

Using the detected N_l anatomical landmarks, scaled rigid registrations are performed to align all the N_A atlases to the target image space. Then we select N_s most similar atlases for the given target image. This is achieved by comparing the sum of the distance of the landmarks for all the aligned atlases after the scaled rigid registration . The selected atlases are further registered to the target image with a Markov Random Field (MRF) based non-rigid registration [18]. We then use the selected atlases to generate probabilistic atlas (PA) for pelvis, bilateral proximal femurs and background following the idea introduced in [19]. The generated PAs are further incorporated to a Maximum-a-Posteriori (MAP) estimation which is then optimized by a graph cut method [20] to obtain segmentation results.

3 Graph optimization based hip joint surface detection

3.1 Problem formulation

After we extract surface models of the pelvis and femur using multi-atlas based segmentation, we expect to refine the hip joint segmentation in the second stage by separating two surfaces of the adjacent structures, i.e., separating the surface of the acetabulum from the surface of the femoral head. In the CT image space, both the acetabulum and the femoral head are ball-like structures and their surfaces can be approximately represented as half-spherically shaped models. To separate these two surfaces, directly applying graph optimization-based surface detection in the CT image space as described in [10, 12] would be an option. However, construction of a graph in the original CT image is not straightforward and requires finding correspondences between two adjacent surfaces obtained from a rough segmentation stage as done in [10, 12], which is challenging.

In our method, instead of performing surface detection in the original CT image space, we first define a hip joint area in the CT image based on the multi-atlas-based segmentation results, and then unfold this area using a spherical coordinate transform as shown in Fig.4. Since the spherical coordinate transform converts a half-spherically shaped surface to a planar surface, the surfaces of the acetabulum and the femoral head can therefore be unfolded to two terrain-like surfaces with a gap (joint space) between them as shown in Fig.4. We reach this goal with following steps:

1. Detecting rim points of the acetabulum from segmented surface model of the pelvis using the method that we developed before [21] (Fig.4: 1).
2. Fitting a circle to the detected rim points, determining radius R_c and center of the circle, as well as normal to the plane where the fitted circle is located (Fig.4: 2).

3. Constructing a spherical coordinate system as shown in Fig.4: 3, taking the center of the fitted circle as the origin, the normal to the fitted circle as the fixed zenith direction, and one randomly selected direction on the plane where the fitted circle is located as the reference direction on that plane. Now, the position of a point in this coordinate system is specified by three numbers: the radial distance R of that point from the origin, its polar angle Θ measured from the zenith direction and the azimuth angle Φ measured from the reference direction on the plane where the fitted circle is located.
4. Sampling points in the spherical coordinate system from the hip joint area (see Fig.4: 4) using a radial resolution of 0.25 mm and angular resolutions of 0.03 radians (for both polar and azimuth angles). Furthermore, we require the sampled points satisfying following conditions:

$$\begin{cases} R_c + 10 \leq R \leq R_c/2 \\ 0 \leq \Theta \leq \pi/2 \\ 0 \leq \Phi \leq 2\pi \end{cases} \tag{3}$$

5. Getting corresponding intensity values of the sampled points from the CT image, which finally forms an image volume $I(\theta, \varphi, r)$ (Fig.4: 5), where $0 \leq r \leq \lceil (10 + \frac{R_c}{2})/0.25 \rceil$, $0 \leq \theta \leq \lceil \pi/0.06 \rceil$ and $0 \leq \varphi \leq \lceil 2\pi/0.03 \rceil$. The dimension of r depends on the radius of the fitted circle while the dimensions of θ and φ are fixed. To easy the description later, here we define the dimension of r as D_r.

Fig.5 shows an example of the unfolded volume $I(\theta, \varphi, r)$ of a hip joint. With such an unfolded volume, graph construction and optimal multiple-surface detection will be straightforward when the graph optimization-based multiple-surface detection strategy as introduced in [11, 12] is used.

3.2 Graph construction for multi-surface detection

For the generated volume $I(\theta, \varphi, r)$ as shown in Fig.5, we assume that r is implicitly represented by a pair of (θ, φ), e.g. $r = p(\theta, \varphi)$. For a fixed (θ, φ) pair, the voxel subset $\{I(\theta, \varphi, r) | 0 \leq r < D_r\}$ forms a column along the $r-axis$ and is defined as $Col(p)$. Each column has a set of neighbors and in this paper 4-neighbor system is adopted. The problem is now to find k coupled surfaces such that each surface intersects each column exactly at one voxel. In our case, we expect to detect two adjacent surfaces of a hip joint, i.e., the surface of the acetabulum S_a and the surface of the femoral head S_f. To accurately detect these two surfaces using graph optimization-based approach, following geometric constraints need to be considered:

1. For each individual surface, the shape changes of this surface on two neighboring columns $Col(p)$ and $Col(q)$ are constrained by smoothness conditions. Specifically, if $Col(p)$ and $Col(q)$ are neighbored columns along the $\theta - axis$, for each surface S (either S_a or S_f), the shape change should satisfy the constraint of $|S(p) - S(q)| = |r_p - r_q| \leq \Delta_\theta$, where $r_p = p(\theta_1, \varphi)$ and $r_q = q(\theta_2, \varphi)$

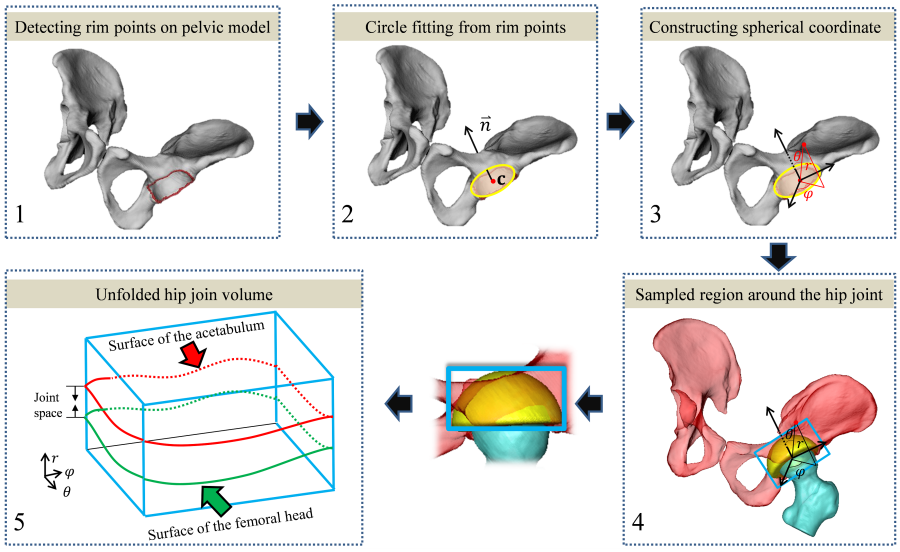

Fig. 4. A schematic illustration of defining and unfolding a hip joint. Please see text in Section 3.1 for a detailed explanation.

are coordinate values of the surface S (either S_a or S_f) intersecting columns $Col(p)$ and $Col(q)$, respectively. The same constraint could also be applied along the $\varphi - axis$ with a smoothness parameter Δ_φ.

2. For the pair of surfaces S_a and S_f, their surface distance in the same column is constrained. For example, in column $Col(p)$, the distance between these two surface should be constrained in a specified range of $0 \leq \delta_p^l \leq (|S_a(p) - S_f(p)|) \leq \delta_p^u$. In addition, S_f requires to be located below the S_a (as shown in Fig.5).

To enforce above geometric constraints, three types of arcs are constructed to define a directed graph $G = \{G_a \cup G_s\}$ (see Fig.6 for details), where G_a and G_s are two subgraphs and each for detecting one surface of S_a and S_f, respectively. For each subgraph, we construct both *intra-* and *inter-* column arcs. We also construct *inter*-surface arcs between two subgraphs G_a and G_s, following the graph construction methods introduced in [11, 12].

Intra-column arcs: This type of arcs is added to ensure that the target surface intersects each column at exactly one position. In our case, along each column $p(\theta, \varphi)$, every node $V(\theta, \varphi, r)$ has a directed arc to the node immediately below it $V(\theta, \varphi, r-1)$ with $+\infty$ weight (Fig.6 left).

Inter-column arcs: This type of arcs is added to constrain the shape changes of each individual surface S on neighboring columns under a 4-neighborhood system. With two pre-defined smoothness parameters Δ_θ and Δ_φ, we construct these arcs with $+\infty$ weight along both the $\theta - axis$ and the $\varphi - axis$ (Fig.6 left).

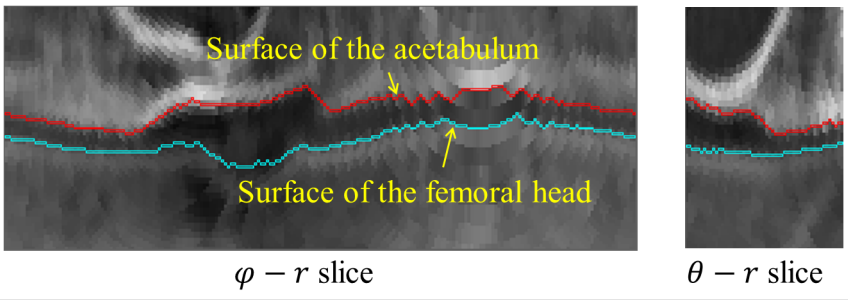

Fig. 5. An example of unfolded volume $I(\theta, \varphi, r)$ of a hip joint, visualized in 2D slices. **Left**: a 2-D φ-r slice. **Right**: a 2-D θ-r slice. In both slices, the green line indicates the surface of the femoral head and the red line indicates the surface of the acetabulum. The gap between these two surface corresponds to the joint space of the hip.

In summary, we have arcs:

$$E = \begin{cases} \{< V(\theta, \varphi, r), V(\theta + 1, \varphi, max(0, r - \Delta_\theta)) >\} \cup \\ \{< V(\theta, \varphi, r), V(\theta - 1, \varphi, max(0, r - \Delta_\theta)) >\} \cup \\ \{< V(\theta, \varphi, r), V(\theta, \varphi + 1, max(0, r - \Delta_\varphi)) >\} \cup \\ \{< V(\theta, \varphi, r), V(\theta, \varphi - 1, max(0, r - \Delta_\varphi)) >\} \end{cases} \tag{4}$$

To get a smooth segmentation, we further enforce soft smoothness shape compliance by adding another type of intra-column arcs (Fig.6 left) [12]:

$$E = \begin{cases} \{< V(\theta, \varphi, r), V(\theta + 1, \varphi, r) > | r \geq 1\} \cup \\ \{< V(\theta, \varphi, r), V(\theta - 1, \varphi, r) > | r \geq 1\} \cup \\ \{< V(\theta, \varphi, r), V(\theta, \varphi + 1, r) > | r \geq 1\} \cup \\ \{< V(\theta, \varphi, r), V(\theta, \varphi - 1, r) > | r \geq 1\} \end{cases} \tag{5}$$

Again we construct these arcs along both the $\theta - axis$ and the $\varphi - axis$ using a 4-neighborhood system. The a *prior* shape compliance smoothness energy that assigned to these arcs are determined by a non-decreasing function $f_{p,q}(|S(p) - S(q)|)$, where $|S(p) - S(q)|$ represent the shape change (determined by the smoothness parameters Δ_θ and Δ_φ) for a surface S on neighbored columns $Col(p)$ and $Col(q)$. We select a linear function $f_{p,q}(|S(p) - S(q)|) = a(|S(p) - S(q)|) + b$, following the method introduced in [12]. Thus, along the $\theta - axis$, we assign a weight a to the arcs. Likewise, for the arcs along the $\varphi - axis$, we assign a similar weight for each arc.

Inter-surface arcs: This type of arcs is added to constrain surface distance between S_a and S_f in each column. In our case S_f is required to be below the S_a. Thus, assuming that distance in column p between surfaces S_a and S_f ranges

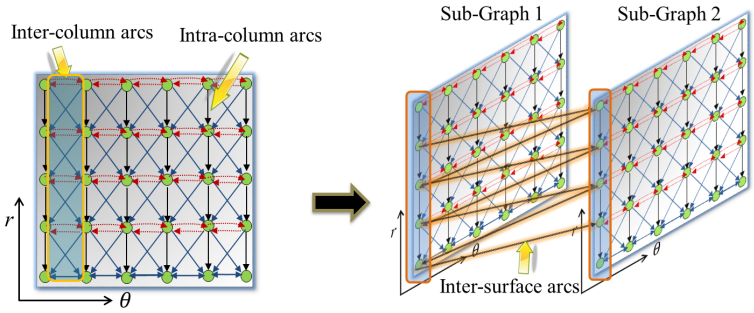

Fig. 6. Graph construction for detecting adjacent two surfaces of a hip joint. An example is presented in 2-D r-θ slice from the unfolded volume $I(\theta, \varphi, r)$. Left: intra-column (black arrows) and inter-column(red and blue) arcs for each subgraph; right: inter-surface arcs to connect two subgraphs. Please note that these two subgraphs share the same nodes as well as the same inter- and intra-column arcs. The inter-surface arcs are constructed between the corresponding two columns which have exactly the same column of voxels in the unfolded volume.

from δ_p^l to δ_p^u, we add the following arcs (Fig.6 right):

$$E_s = \begin{cases} \{< V_a(\theta, \varphi, r), V_f(\theta, \varphi, r - \delta_p^u) > | r \geq \delta_p^u\} \cup \\ \{< V_f(\theta, \varphi, r), V_a(\theta, \varphi, r + \delta_p^l) > | r < R - \delta_p^l\} \cup \\ \{< V_a(0, 0, \delta_p^l), V_f(0, 0, 0) >\} \end{cases} \tag{6}$$

where V_a and V_f denote the node in the corresponding column from each subgraph as shown in (Fig.6 right). For each column p, we have a different distance range (δ_p^l, δ_p^u) that is statistically learned from a set of training data.

By adding all the arcs as described above, we establish a directed graph $G = (V, E)$, where $V = V_a \cup V_f$ and $E = E_a \cup E_f \cup E_s$. Here, V_a and V_f are node sets from each subgraph, E_a and E_f are intra- and inter-column arcs from each subgraph, and E_s is the inter-surface arcs between two subgraphs. In order to detect surfaces based on graph optimization, a new digraph $G_{st}(V \cup \{s, t\}, E \cup E_{st})$ is defined. This is achieved by adding a source node s and a sink node t as well as new edge set E_{st} which includes the edges between nodes in the graph G and the nodes of $\{s, t\}$. Then surface detection can be solved using the minimum s-t cuts established by Kolmogorov et al. [22]. We add new edges for the edge set E_{st} following the method introduced in [11]. The most important point here is to assign an appropriate penalty for each edge which is also called $t-link$. As described in [11], the penalty for each $t-link$ is determined by a pre-computed cost of each node. In our method, a carefully designed node cost function is calculated by considering both intensity information and *a prior* information. In the next section, we will introduce how such a cost function is calculated for each graph node.

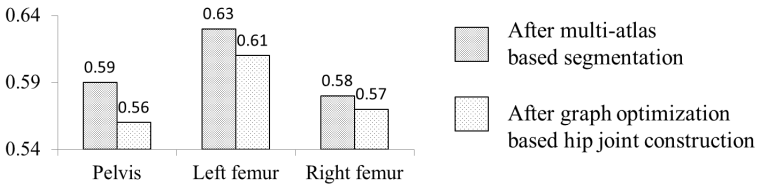

Fig. 7. Evaluation of the segmentation performance after each stage in the present method. The average surface distance (mm) on 10 test images are shown in the figure.

3.3 Node cost function

Node cost function plays an important role for a successful surface detection. In our method, we first encode the boundary information to the cost function using the gradient information of each node following the method introduced in [11]. The negative magnitude of the gradient of the volume $I(\theta, \varphi, r)$ is computed at each voxel as $c_{edge}(\theta, \varphi, r) = -| \bigtriangledown I(\theta, \varphi, r)|$. We give each node a weight as:

$$w(\theta, \varphi, r) = \begin{cases} c_{edge}(\theta, \varphi, r) & if \quad z = 0 \\ c_{edge}(\theta, \varphi, r) - c_{edge}(\theta, \varphi, r-1) & otherwise \end{cases} \tag{7}$$

These weights are then modified by adding three types of constraints from a *prior* information: 1) The generated PA of the pelvis and the femur in the multi-atlas based segmentation stage. 2) Intensity histograms of surface points which approximately indicate the intensity distribution of the points on each surface. They are statistically learned from a set of training data by extracting all the points on manually segmented surfaces from each training data. We learned two histograms, one for the acetabulum surface and the other for the femoral head surface. 3) The orientation of the gradient along the $r-axis$, which is determined by a sign function $Sgn(I_i(\theta, \varphi, r) - I_i(\theta, \varphi, r-1))$, where $I_i(\theta, \varphi, r)$ is the intensity of voxel $V(\theta, \varphi, r)$.

The PA gives the probability of each voxel belonging to a specified bone region (background, pelvis or bilateral proximal femurs). If any voxel have probability close or equal to 1 in the PA, it means that the atlases used for generating PA voted for this voxel, and thus this voxel is more likely to appear inside the bone region rather than on the surface of the bone. Considering that our purpose is to detect the surfaces of the bones, we decrease the weights for such nodes by:

$$w'(\theta, \varphi, r) = \begin{cases} w(\theta, \varphi, r) - h \cdot PA(\theta, \varphi, r), & if \quad w(\theta, \varphi, r) > 0 \\ w(\theta, \varphi, r) + h \cdot PA(\theta, \varphi, r), & if \quad w(\theta, \varphi, r) \leq 0 \end{cases} \tag{8}$$

where h is a constant value and $PA(\theta, \varphi, r)$ is the probability for a voxel $V(\theta, \varphi, r)$. For the nodes in the subgraph for detecting the surface of the acetabulum, we perform such a modification using the PA of the pelvis. Likewise, we encode information from the PA of the femur for the nodes in the other subgraph.

Table 1. Surface distance (mm) between automatic and ground-truth segmentation of the bilateral hip joints from 10 CT data. Results after stage I (multi-atlas-based segmentation: MA) and after stage II (graph optimization-based surface detection: GO) are shown, where LA stands for the left acetabulum, LFH for the left femoral head, RA for the right acetabulum and RFH for the right femoral head.

Bone	Stage	CT 1	CT 2	CT 3	CT 4	CT 5	CT 6	CT 7	CT 8	CT 9	CT 10	Average
LA	MA	0.42	0.30	0.24	0.24	0.26	0.30	0.46	0.40	0.29	0.35	0.33
	GO	0.20	0.29	0.21	0.23	0.23	0.17	0.31	0.26	0.26	0.24	0.24
LFH	MA	0.40	0.56	0.41	0.34	0.51	0.36	0.51	0.81	0.38	0.49	0.48
	GO	0.30	0.25	0.24	0.37	0.33	0.32	0.41	0.64	0.31	0.27	0.34
RA	MA	0.43	0.36	0.22	0.33	0.24	0.30	0.48	0.31	0.25	0.32	0.32
	GO	0.23	0.23	0.18	0.30	0.17	0.22	0.22	0.21	0.22	0.30	0.23
RFH	MA	0.40	0.42	0.54	0.29	0.34	0.39	0.55	0.52	0.43	0.53	0.44
	GO	0.41	0.40	0.46	0.36	0.38	0.38	0.45	0.59	0.38	0.28	0.41

The intensity distribution of the bone surface points is limited in a specified range. For voxels whose intensity values are in this specified range, they are more likely to appear on the bone surfaces. We increase the weights for these nodes using the associated intensity histograms of surface points that are learned from a set of training data as described above.

$$w''(\theta,\varphi,r) = \begin{cases} w'(\theta,\varphi,r) + m \cdot Hist(I_i(\theta,\varphi,r)), & if \;\; w(\theta,\varphi,r) > 0 \\ w'(\theta,\varphi,r) - m \cdot Hist(I_i(\theta,\varphi,r)), & if \;\; w(\theta,\varphi,r) \leq 0 \end{cases} \tag{9}$$

where m is a constant value and $Hist(I_i(\theta,\varphi,r))$ is the corresponding value in the associated histogram for voxel $V(\theta,\varphi,r)$ which have intensity value of $I_i(\theta,\varphi,r)$. Please note that we have learned two intensity histograms, one for surface points of the acetabulum and the other for the surface points of the femoral head. When modifying the weights of nodes in each subgraph, the associated histogram is used.

As shown in Fig.5, for voxels on the surface of the acetabulum, since the intensity at $V(\theta,\varphi,r)$ is bigger than $V(\theta,\varphi,r-1)$), the value of the $Sgn(I_i(\theta,\varphi,r) - I_i(\theta,\varphi,r-1))$ should be positive and we define the orientation of these voxels as positive too. Similarly, for voxels on the surface of the femoral head, we define their orientation in $r-axis$ as negative. Therefore, for a node in the subgraph for detecting the surface of the acetabulum, if its gradient orientation is not consistent with our definition, we set its weight to 0. For a node in the subgraph for detecting the surface of the femoral head, we perform a similar modification.

After we modify the weight for each node in the Graph G_{st}, we assign penalty for each t-link based on the modified weight using the method introduced in[11]. Our problem is then to optimally detect two surfaces from the constructed graph, which can be solved using the minimum s-t cuts algorithm [20, 22].

Table 2. Comparison of the results achieved by the present method with those reported in the literature.

Method	Preserving Hip Joint	Average SD (mm)	Average DOC (%)
Lamecker et al. [9]	No	1.80	-
Semi et al. [4]	No	0.70	-
Kainmueller et al. [5]	Yes	0.60	-
Yokota et al. [2]	No	1.10	92.7
Yokota et al. [3]	No	0.98	-
The present method	Yes	0.58	94.7

4 Experiments and results

We evaluated the present method on hip CT data of 30 patients after ethical approval. The intra-slice resolutions range from 0.576mm to 0.744mm while the inter-slice resolutions are 1.6mm for all CT data. Manual segmentation of all 30 CT data were done by a trained rater. 20 of them were selected as the training data both for the RF regression-based landmark detection and the multi-atlas-based segmentation. The rest 10 datasets (20 hip joints) were used for evaluation.

As for performance evaluation, we computed two different metrics. First, surface distance (SD) between automatic segmentation and ground-truth segmentation are computed after each stage of the present method. Fig. 7 shows the average SD which was computed on the entire pelvis and femur regions. More specifically, our method achieves high performance in segmenting pelvis, left proximal femur, and right proximal femur with an average accuracy of 0.56 mm, 0.61 mm, and 0.57 mm, respectively. Furthermore, we also looked at the segmentation accuracy around the hip joint local areas which are important for our target clinical applications. The local evaluation results are shown in Table 1. It is observed from Fig.7 that the segmentation results of the two stages in the present method are quite close if we evaluate the performance in the entire regions of the pelvis and the femur. However, when we focus on the local hip joint area, one can find from Table 1 that the graph optimization-based surface detection improves the hip joint segmentation accuracy. Second, we also computed the Dice overlap coefficients (DOC) between automatic segmentation and ground-truth segmentation. The present method achieved a mean DOC of 93.3±1.1%, 95.2±1.3%, and 95.5±0.8% for pelvis, left femur and right femur, respectively.

We checked whether the present method could preserve the hip joint space and prevent the penetration of the extracted surface models. For all the 20 hip joints that were segmented with the present method, we have consistently found that the hip joint spaces were preserved and that there was no penetration between the extracted adjacent surface models. Fig. 8 shows a segmentation example. From this figure, we can clearly see that the graph optimization-based surface detection stage further improve the results from the multi-atlas-based segmentation.

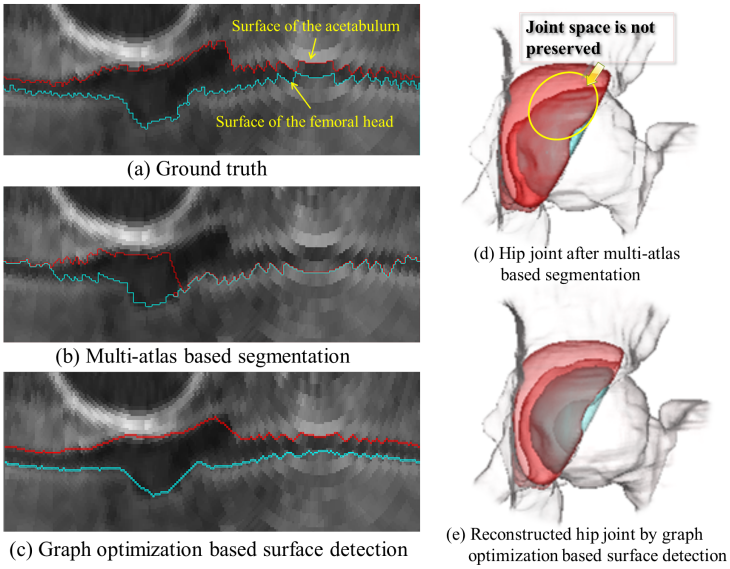

Fig. 8. An example of segmenting a hip joint with the present method. Both 2D and 3D visualizations of results from different stages of our method are presented.

5 Discussions and Conclusion

The goal of the present study is to develop and validate a fully automatic hip joint segmentation approach. Our experimental results showed that the present method not only achieved a good overall segmentation accuracy for both the pelvis and the proximal femur, but also had the advantages of preservation of hip joint space and prevention of the penetration of the extracted adjacent surface models, which are prerequisite conditions to use the segmented models for computer assisted diagnosis and planning of PAO surgery.

The performance of the present method is compared with those of the state-of-the-art hip CT segmentation methods [2–5, 9]. The comparison results are summarized in Table 2. From this table, one can see that the performace of the present method is comparable to other state-of-the-art hip CT segmentation methods [2–5, 9].

In conclusion, we presented a fully automatic and accurate method for segmenting CT images of a hip joint. The strength of the present method lies in the combination of a multi-atlas-based hip CT segmentation with a graph optimization-based multi-surface detection. The present method can be extended to segment CT data of other anatomical structures.

References

1. Ganz, R., Klaue, K., Vinh, T., Mast, J.: A new periacetabular osteotomy for the treatment of hip dysplasia. technique and preliminary results. Clin Orthop **232** (1988) 26–36
2. Yokota, F., Okada, T., Takao, M., Sugano, S., Tada, Y., Sato, Y.: Automated segmentation of the femur and pelvis from 3D CT data of diseased hip using hierarchical statistical shape model of joint structure. In Yang, G., Hawkes, D., Rueckert, D., Noble, A., Taylor, C., eds.: MICCAI 2009. Volume 5762., LNCS (2009) 811–818
3. Yokota, F., Okada, T., Takao, M., Sugano, S., Tada, Y., Tomiyama, Y., Sato, Y.: Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models. In Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N., eds.: MICCAI 2013. Volume 8150., LNCS (2013) 190–197
4. Seim, H., Kainmueller, D., Heller, M., Lamecker, H., Zachow, S., Hege, H.C.: Automatic segmentation of the pelvic bones from CT data based on a statistical shape model. In: Eurographics Workshop on Visual Computing for Biomedicine. (2008) 67–78
5. Kainmueller, D., Lamecker, H., Zachow, S., Hege, H.C.: An articulated statistical shape model for accurate hip joint segmentation. In: IEEE EMBC 2009. (2009) 6345–6351
6. Ehrhardte, J., Handels, H., Plotz, W., Poppl, S.J.: Atlas-based recognition of anatomical structures and landmarks and the automatic computation of orthopedic parameters. Methods Inf. Med **43** (2004) 391397
7. Pettersson, J., Knutsson, H., Borga., M.: Automatic hip bone segmentation using non-rigid registration. In: ICPR 2006. (2006)
8. Ying, X., Jurgen, F., Shekhar, S., Raphael, S., Craig, E., Stuart, C.: Automated bone segmentation from large field of view 3D MR images of the hip joint. Physic in Medicine and Biology **58** (2013) 7375–7390
9. Lamecker, H., Seeba, M., Hege, H.C., Deuflhard, P.: A 3D statistical shape model of the pelvic bone for segmentation. In: SPIE. Volume 5370. (2004) 1341–1351
10. Kainmueller, D., Lamecker, H., Zachow, S., Hege, H.C.: Coupling deformable models for multi-object segmentation. In: ISBMS. Volume 5104. (2008) 6978
11. Li, K., Wu, X., Chen, D., Sonka, M.: Optimal surface segmentation in volumetric images - a graph-theoretic approach. IEEE Trans. Pattern Anal. Machine Itell **28** (2006) 119–134
12. Song, Q., Wu, X., Liu, Y., Smith, M., Buatti, J., Sonka, M.: Optimal graph search segmentation using arc-weighted graph for simultaneous surface detection of bladder and prostate. In: MICCAI 2009. Volume 5762., LNCS (2009) 827–835
13. Criminisi, A., Shotton, J., Robertson, D., Konukoglu, E.: Regression forests for efficient anatomy detection and localization in ct studies. In Menze, B., Langs, G., Tu, Z., Criminisi, A., eds.: MICCAI 2011. Volume 6533., LNCS (2011) 106–117
14. Lindner, C., Thiagarajah, S., Wilkinson, J.M., arcOGEN Consortium, Wallis, G., Cootes, T.F.: Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE TMI **32** (2013) 1462–1472
15. Breiman, L.: Random forests. Machine Learning **45** (2001) 5–32
16. Yang, C., Duraiswami, R., Davis, L.: Efficient kernel machines using the improved fast gauss transform. Advances in neural information processing systems **17** (2005) 1561–1568

17. P. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: CVPR 2001. Volume I. (2001) 511–518
18. Glocker, B., Komodakis, M., Tziritas, G., Navab, N., Paragios, N.: Dense image registration through mrfs and efficient linear programming. Medical Image Analysis **12** (2008) 731–741
19. Chu, C., Oda, M., Kitasaka, T., Misawa, K., Fujiwara, M., Hayashi, Y., Nimura, Y., Rueckert, D., Mori, K.: Multi-organ segmentation based on spatially-divided probabilistic atlas from 3D abdominal CT images. In Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N., eds.: MICCAI 2013. Volume 8150., LNCS (2013) 165–172
20. Boykov, Y., Veksler, O., Zabih, R.: Fast approximate energy minimization via graph cuts. IEEE PAMI **23** (2001) 1222–1239
21. Liu, L., Ecker, T., Schumann, S., Siebenrock, K., Nolte, L., Zheng, G.: Computer assisted planning and navigation of periacetabular osteotomy with range of motion optimization. In Golland, P., et al., eds.: MICCAI 2014. Volume Part II, LNCS 8674., Springer (2014) 643 – 650
22. Kolmogorov, V., Zabih, R.: What energy functions can be minimized via graph cuts? IEEE Trans. on PAMI **26** (2004) 147–159