

Graph-Cut-Based Segmentation of Proximal Femur from Computed Tomography Images with Shape Prior

Junbin Huang^{1,2} · James F. Griffith¹ · Defeng Wang^{1,2} · Lin Shi^{3,4}

Received: 18 November 2014 / Accepted: 9 April 2015 / Published online: 6 October 2015
© Taiwanese Society of Biomedical Engineering 2015

Abstract Femur segmentation from computed tomography (CT) images is a fundamental problem in femur-related computer-assisted diagnosis and surgical planning/navigation. In this study, an automatic approach for the segmentation of proximal femur from CT images that incorporates the statistical shape prior into the graph-cut framework (SP-GC) is proposed. The proposed segmentation framework includes two major processes, namely training and segmentation. In the training stage, a training set of three-dimensional CT images from a group of normal subjects were segmented semi-automatically. The shape prior was generated by active shape modeling that included mean shape and shape variance information. In the segmentation stage, two shape terms originated from the training stage were included in the GC energy function. The minimization of the energy function was achieved using a max-flow/min-cut algorithm. The performance of the proposed segmentation method was evaluated by testing on 60 CT datasets from bilateral femurs of 30 normal subjects. Qualitative and quantitative analyses of the segmentation results of the proposed method were performed

and the results were compared with two widely used methods, namely the active shape model (ASM) and traditional GC, with results from manual delineation used as the ground truth. The mean dice similarity coefficient of the proposed SP-GC was 0.9600, which is higher than those of ASM and GC (0.8769 and 0.9358, respectively). The mean normalized error rate of the SP-GC results was 10 and 6 % lower than those of ASM and GC, respectively. In terms of the average surface distance measurement, the value for SP-GC was 0.885 mm, compared with 2.148 and 1.154 mm for ASM and GC, respectively. In comparison with ASM, SP-GC has superior performance given a small training set (e.g., $n = 12$). With increasing number of training samples, the segmentation accuracy of ASM saturated, but that of SP-GC slightly increased.

Keywords Graph-cut algorithm · Active shape model · Shape prior · Computed tomography · Proximal femur

1 Introduction

Femur segmentation is a fundamental problem in femur-related research and clinical applications. The detection of the imaging biomarker of femoral head shape in femoroacetabular impingement (FAI) has attracted increasing attention due to its great potential in clinical diagnosis. FAI is a relatively recently defined femur condition that could be a pre-condition of osteoarthritis [1, 2]. Cam-type FAI can be differentiated from pincer-type FAI by a remarkable protrusion on the femoral head-and-neck junction. Clinically, radiographic images or two-dimensional (2D) slices from computed tomography (CT) or magnetic resonance (MR) scans are frequently utilized in its diagnosis [3–5]. However, the reproducibility of 2D-

✉ Lin Shi
shilin@cuhk.edu.hk

¹ Department of Imaging and Interventional Radiology, The Chinese University of Hong Kong, Hong Kong SAR 999077, China

² Research Center for Medical Image Computing, The Chinese University of Hong Kong, Hong Kong SAR 999077, China

³ Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong SAR 999077, China

⁴ Chow Yuk Ho Technology Centre for Innovative Medicine, The Chinese University of Hong Kong, Hong Kong SAR 999077, China

image-based diagnosis is problematic due to the limited shape information presented. In contrast, the three-dimensional (3D) surface of a femur offers much more complete shape information, so that the statistical femoral shape difference can be detected for differentiating cam-type FAI patients from normal subjects and patients with other types of FAI [6]. CT is the preferred imaging modality for capturing 3D bone structure due to its distinguishable high intensity in cortical bone. Therefore, the segmentation of the proximal femoral bone based on CT is suitable for differentiating cam-type FAI patients.

Despite the highly differentiable signal intensity in cortical bone in CT images, automatic segmentation of the femur is still a non-trivial problem. A variety of segmentation algorithms have been proposed [7]. The simplest and fastest method is threshold-based, but is rarely used independently as it is susceptible to noise [8]. This type of segmentation method is successful only when the structure to be segmented (i.e., bone) has intensities that are obviously larger or smaller than those of the background tissues. Therefore, it is seldom possible to determine a threshold value for all the voxels of the target to be larger or smaller than this threshold value that. Some efforts have been made to improve performance, such as multi-threshold [9] and local adaptive threshold methods [10]. Algorithms such as region growing and watershed have the same limitations as those of threshold-based methods and often provide initial segmentation results that need post-processing to be accurate [11]. They tend to cause over-segmentation and cannot find boundaries of significant areas when they are low contrast compared to the background. Recently, the graph-cut (GC) algorithm [12] has shown great promise for medical image segmentation and has thus attracted broad attention. Similar to active contour [13], level set [14], and live wire methods [15], GC is an energy-based method. It converts an image to a graph and transforms the image segmentation problem to an energy function minimization problem. GC combines the information of image intensities and boundaries into its energy function and assumes that when the image is well segmented, this function will reach its minimum. The segmentation methods mentioned above are usually applied directly to raw image data and make no use of any shape priors of the segmentation target. For using shape priors to aid segmentation, the active shape model (ASM) and active appearance model [16, 17] were proposed. The procedure is divided into two stages, namely training and testing. Firstly, images for training are segmented manually in a slice-by-slice manner. Then, the segmentation results are grouped and principal component analysis is used to obtain the mean surface and variability in the group captured by a number of principal modes. The mean surface is a global representation of the training set while the modes give

dominant distinctions among individual samples. In addition to shape variance, the gray-level appearance is generated according to local intensity feature of landmarks. In the testing stage, the ASM and active appearance model search for the updated positions of landmarks by comparing the gray level appearance of current landmarks and their neighboring voxels.

Extraction of the femur from CT images for analysis of the proximal femur, which is of vital importance for characterizing FAI femur morphometry, is the pre-condition for constructing the femoral statistical model. It is difficult to obtain reliable segmentation results using only thresholding, region growing, or other methods. To overcome the limitations of these general methods, priors about the segmentation target can be utilized. For instance, Yokota et al. [18] proposed an automated CT segmentation strategy for abnormal hips using hierarchical and conditional statistical shape models. The method firstly segments the pelvis and femur simultaneously and then refines the segmentation of the femur to improve accuracy under the condition that the distal femur and pelvis were segmented. However, due to the repeated optimizations of the statistical shape model, the computation time is considerable. Although the ASM or active appearance model yields acceptable performance in osseous structure segmentation with CT images [19], the result may be suboptimal but not a global optimum. Moreover, to guarantee the reliability and accuracy of the statistical model, a considerable number of training samples are usually necessary, as the testing stage relies on the training results. In contrast, the GC method uses the max-flow/min-cut algorithm to optimize its energy function and the result is guaranteed to be the global optimum [20]. For example, Krcah et al. [21] used GC to automatically segment the femur bone from 3D CT images with a success rate of 81 % based on 197 samples. In this method, a bone boundary enhancement filter is applied, but the separation of neighboring bones was successful in only 57 % of the cases, which seems unacceptable in practical applications.

Using image features alone (i.e., neglecting the shape prior of segmentation target) is likely to achieve incorrect segmentation results [22]. For instance, the narrow space between the femoral head and the acetabulum is difficult to distinguish with GC. In this situation, parts of the acetabulum are regarded as a femoral component and segmented. Instead of segmenting the proximal femur in a single-step pattern, some methods addressed this issue in a coarse-to-fine manner. In this framework, a threshold-based method is first used regardless of the connection of the femoral head and the acetabulum. The main effort is then taken to separate the femoral head and the acetabulum as these two structures should be disconnected actually. Zoroofi et al. [23] used Otsu's method to first segment bone tissues, after

which the initial boundary of the femoral head and the acetabulum as well as the joint space between them were estimated utilizing the information of the greater trochanter and the shape of the femoral head. The narrow space between the femur and the acetabulum was detected using the moving disk technique. This technique achieved good segmentation accuracy in about 70 % of cases. Cheng et al. [24] used Bayes decision rules to consider the neighborhood information and the partial volume effect to detach a connected femur joint after histogram-based thresholding. Based on a deformable model, O'Neill et al. [25] used the morphological snake algorithm to iteratively extract cam-type femurs on a slice-by-slice basis after dividing the femur into the femoral head and the body. However, as a semi-automatic framework, necessary interaction is needed and for the concave region between the femoral neck and the greater trochanter, segmentation performance is qualitatively unacceptable.

Methods combining the original GC with other energy terms have been proposed. Nakagomi et al. [26] presented a GC algorithm with neighbor prior constrains for lung segmentation in CT images. They introduce priors on neighboring structures from a probabilistic atlas and perform the segmentation in a coarse-to-fine manner. Another method uses weighted directed graph construction to impose geometrical and smooth constraints learned from priors and built a cost function by combining selective feature extractors using a support vector machine classifier [27]. Level set has been used for prostate segmentation in MR images [28]. The GC results are used as the initial values for the level set, after which level-set-based segmentation is used with a shape prior. In this method, level set is the main algorithm. In GC segmentation with an adaptive shape prior, instead of using a shape prior at all pixels, the prior is obtained from selected pixels, where image labels are difficult to determine to overcome the parameter selection problem [29]. Chen et al. [30] proposed a synergistic combination of GC and ASM for medical image segmentation by adding a shape term to provide additional information for the data term in the original GC. An overall accuracy of 96 % was achieved with the method in 2D image segmentation. Intuitively, it is reasonable to incorporate a shape prior as extra information for the boundary term. In the present study, the different influences of shape prior as supplementary data and boundary information are investigated.

The proposed segmentation scheme of the proximal femur using CT images is based on the GC algorithm incorporating a shape prior (SP-GC). A point distribution model is used for the training samples and the variances of each landmark among the training group is generated. Then, these statistical variant modes are used to define the shape term in the GC energy function. For landmark

correspondence, a plane that passes the most proximal point of the lesser trochanter and is perpendicular to the main axis of the femoral shaft for each sample is first determined interactively. The components above these planes are registered among training samples, after which a slightly higher plane is defined for the whole dataset to obtain training shapes for shape-context-based correspondence [31]. The shape term is attached to the original GC energy function and the max-flow/min-cut algorithm is used to obtain the final segmentation results. The segmentation accuracy was measured using the dice similarity coefficient (DSC), normalized error rate [32], and average shape distance to compare the performances of the SP-GC, ASM, and original GC algorithms. Furthermore, for SP-GC and ASM, the influence of the training set scale on segmentation reliability was evaluated for various numbers of training samples (12, 16, 20, 24, 28, 32, 36 and 40).

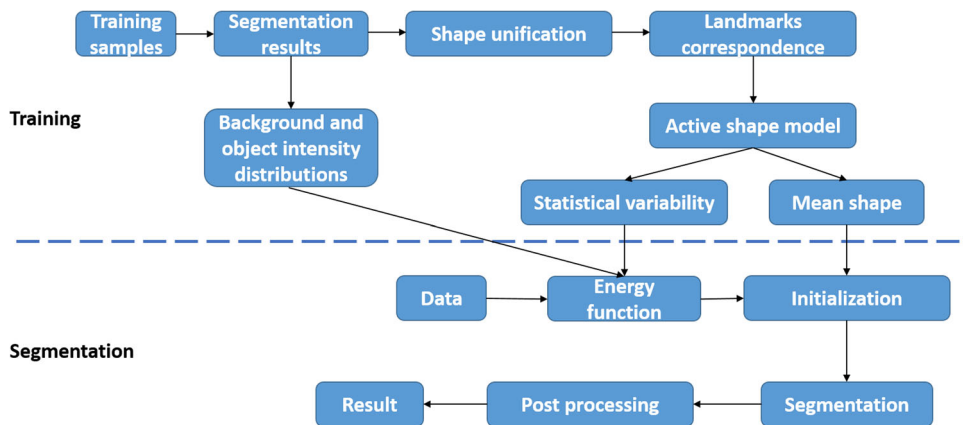
2 Materials and Methods

The conceptual framework of the proposed segmentation method is shown in Fig. 1. The proposed automatic segmentation framework can be divided into two stages, namely the training stage and the segmentation stage. In the training stage, training samples are first segmented. In the segmentation stage, the background and object intensity distributions are obtained from the original images and segmented binary masks for constructing the energy function. Shape unification is used to unify training shapes and reduce correspondence error (see Sect. 2.4.1). Generally, the generated training shapes are described with various numbers of points. Then, landmark correspondence is applied to unify the numbers of vertices as well as produce correspondence across samples. With a point distribution model, the mean shape of the proximal femur and the statistical variability are obtained for the next stage. In the segmentation stage, an energy function is constructed based on the intensity distributions from training samples and ASM results. The input data and intensity distributions are used to define data and boundary terms, as in the original GC, while statistical variances from the training stage are for constructing the new shape term. After initializing the mean shape on input data by setting the translational, rotational and scaling factors, followed by the minimization optimization using max-flow/min-cut algorithm, the final segmentation result will be achieved.

2.1 Data Acquisition

A total of 84 femoral CT scans was used in this study. Among the 84 femurs, 40 samples were selected randomly as candidates for model training and the remaining 44

Fig. 1 Flow chart of SP-GC algorithm



femurs and 16 randomly chosen samples from the training dataset (for a total of 60) were used as the testing subset. Participants were selected from those who accepted CT examination of the hip. The CT scans covered the range between the whole proximal femur and the lesser trochanter. The inclusion criteria were that the subject should be asymptotic of any femoral diseases and that the appearance of the femur should be normal. CT examination was performed on a 64-slice multi-detector CT scanner (LightSpeed Ultra, GE Healthcare, Milwaukee, WI, USA) with a 50-cm scan field of view, 0.625-mm-thick slice acquisition with no spacing, an accelerating voltage of 120 kVp, an X-ray tube current of 300 mA, and a 0.6-s scan time. Image reconstruction was conducted using a standard reconstruction algorithm for bone.

2.2 Graph-Cut Algorithm

Image segmentation aims to divide an image into several partitions based on the distinctions of some image features, such as intensity and texture [33]. Image segmentation is analogous to cutting a graph in graph theory. To apply GC to medical image segmentation, the image is first converted into a graph. Based on image information, vertices and edges of the constructed graph are weighted. Image segmentation is then transformed into a numerical energy minimization problem.

To apply the GC algorithm, an image is transformed into an S-T graph $G = \langle V, E \rangle$, where V is the set of vertices and E is the set of edges of an undirected graph G . In GC, two kinds of vertices as well as edges are defined. Each voxel of the image corresponds to a vertex in G . Another type of vertex, called terminal vertices, are virtual nodes used to represent the background and object. There are only two terminal vertices, namely S for the object (source) and T for the background (sink). The edges constructed to connect two neighboring voxels are called t-edges while the lines linking each voxel to the terminal vertices (S or T)

are n-edges [34]. In Fig. 2, a 2D image with 9 pixels is converted into an S-T graph with 11 vertices (9 vertices corresponding to pixels and two terminal vertices). According to the weights allocated to each vertex and edge, the GC algorithm divides the image into the background and the object (red curve).

The GC algorithm solves the energy minimization problem using the max-flow/min-cut algorithm. For image segmentation, the energy function tries to embed image information into the weights of vertices and edges in the graph. The cost function in GC consists of its data term and a boundary term:

$$C(L) = \lambda D(L) + (1 - \lambda)B(L) \tag{1}$$

where $C(L)$ is the cost of the optimal segmentation separating the image into two parts. $D(L)$ and $B(L)$ are the data term and boundary term, respectively. L is the set of labels of image voxels, whose value is either 0 or 1, representing the background and the object, respectively. λ is a user-

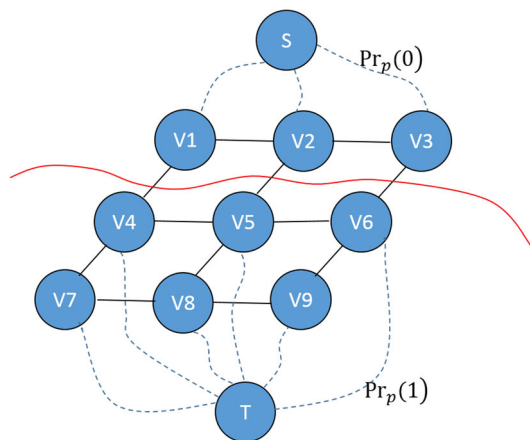


Fig. 2 S-T graph where black solid lines and blue dashed curves represent t-edges and n-edges, respectively. Vertices 1, 2, and 3 belong to the object and remaining 6 vertices belong to the boundary divided by red curve

defined constant to balance the contributions of $D(L)$ and $B(L)$.

The definitions of $D(L)$ and $B(L)$ are based on the segmentation image itself to reflect image features. With the intensity distributions of the background and the object, $D(L)$ can be defined as:

$$D(L) = \sum_{p \in V} Pr_p(L_p) \begin{cases} Pr_p(0) = -\ln Pr(I_p | 'background') \\ Pr_p(1) = -\ln Pr(I_p | 'object') \end{cases} \quad (2)$$

where p is one of the vertices in V and L_p is the label of p . $Pr_p(L_p)$ is commonly defined as the negative logarithm of the probability when assigning L_p to p , which can be computed from its intensity I_p and the distributions of the background and the object (Fig. 2).

The boundary term $B(L)$ is defined based on both n-edges and t-edges to reflect boundary information of the image. The definition of $B(L)$ is as follows:

$$B(L) = \sum_{\langle p_1, p_2 \rangle \in E} \delta(L_{p_1}, L_{p_2}) B_{\langle p_1, p_2 \rangle} \delta(L_{p_1}, L_{p_2}) \\ = \begin{cases} 0, & \text{if } (L_{p_1} = L_{p_2}) \\ 1, & \text{if } (L_{p_1} \neq L_{p_2}) \end{cases} \quad (3)$$

where $\langle p_1, p_2 \rangle$ is an edge of E linking two neighboring vertices p_1 and p_2 . $B_{\langle p_1, p_2 \rangle}$ should be a non-increasing function related to the intensity difference between p_1 and p_2 , which is commonly proportional to the Gaussian function as:

$$B_{\langle p_1, p_2 \rangle} \propto \exp\left(-\frac{(I_{p_1} - I_{p_2})^2}{2\sigma^2}\right) \frac{1}{Dist(p_1, p_2)} \quad (4)$$

where $Dist(p_1, p_2)$ is the Euclidean distance between p_1 and p_2 . σ is a factor to be defined empirically. In this formula, when the intensities of p_1 and p_2 are very similar, $B_{\langle p_1, p_2 \rangle}$ will be very large and it will be practically impossible for a cut to occur on edge $\langle p_1, p_2 \rangle$. With construction of the weights on vertices and edges, the energy function of GC can be minimized using the max-flow/min-cut algorithm. The result is then transformed back into an image and the segmentation can be described as a binary mask.

2.3 Active Shape Modeling

ASM is method for deriving a statistical model of a group of shapes [16]. Frequently, the input of the ASM framework is a certain number of pre-segmented data that are presented as surfaces with vertices and triangulations using the marching cubes algorithm [35]. ASM aims at generating a mean shape by averaging all the input samples and generalizing variability among individual shapes using

principal component analysis [36]. Usually, the shape description is stacked as a column vector $S_i = (x_1, y_1, z_1, \dots, x_n, y_n, z_n)^T$, where n is the number of landmarks and the triple (x_j, y_j, z_j) , $j = 1, 2, \dots, n$ is the Cartesian coordinates of a landmark. For a training set consisting of t surfaces, the mean surface can be generated with:

$$\bar{S} = \frac{1}{t} \sum_{i=1}^t S_i \quad (5)$$

To generalize individual variability inside the group, singular value decomposition [37] is applied on the covariance matrix:

$$C = \frac{1}{t-1} \sum_{i=1}^t (S_i - \bar{S})(S_i - \bar{S})^T \quad (6)$$

By performing singular value decomposition, the eigenvalues and corresponding eigenvectors of C can be obtained. These eigenvectors delineate how each individual sample in the training group varies in relation to the mean shape and the eigenvalues describe which variability mode is most prominent (the others contribute little to the differences). In this way, each shape used for training can be approximately described as:

$$S_i \approx \bar{S} + \sum_{i=1}^v P_i b_i \quad (7)$$

where b_i is the scalar coefficient of ASM and P_j is the corresponding eigenvectors obtained from singular value decomposition. Generally, not all of the variability modes are used to construct a specific shape. To determine the first v modes used for shape construction, the chosen v should satisfy:

$$\sum_{i=1}^v \lambda_i \geq \alpha_v \sum_{i=1}^{3n} \lambda_i \quad (8)$$

where α_v is a positive parameter and λ_i is the i th eigenvalue when ranking all the eigenvalues from largest to smallest. Usually, α_v will be not smaller than 90 %. As 3D shapes are considered in our experiments, the total number of eigenvalues is $3n$.

2.4 Graph-Cut with Shape Priors

2.4.1 Shape Unification

Accurate manual segmentation results are required for the training stage in our segmentation framework. To convert the segmented 3D binary images into surfaces represented by vertices and triangulations, the marching cubes

algorithm is used. Under normal circumstances, the number of vertices describing each sample should be different when generated, in which situation the point correspondence is necessary. Being different from skull whose landmarks can be defined on obvious biomarkers (e.g. the brow ridges and afterbrain) [19], feature point definition on the femur surface to unify data is more vulnerable to operators as the different scanning ranges for different individuals. After reconstruction, Harris et al. [6] cropped the femur at the superior aspect of the lesser trochanter, which is considered to be the most possible position where cam-tape FAI may occur. However, no more details were given on how to define those cutting planes. In order to obtain a more accurate point correspondence result, a semi-automatic method is applied to unify data for point correspondence. First, the most proximal vertex of the lesser trochanter is found and a plane that passes through this vertex and is perpendicular to the body shaft is approximately defined. The component including the femoral head is then obtained. In addition, the partition of each sample initially defined at this stage is aligned to a randomly chosen template by performing rigid registration with the iterative closest point algorithm [38]. Finally, a new plane parallel to the formal one for acquiring the template and 1 cm closer to the femoral head serves as the unified plane. Components above the unified plane of all the samples are provided as inputs for point correspondence using shape context [31] (Fig. 3).

2.4.2 Shape Prior

The original GC algorithm utilizes intensity and boundary information of the image to perform segmentation. Without the information of the desired segmentation result, the GC algorithm tends to find incorrect boundaries in some situations. The shape prior, a useful piece of information for guiding image segmentation, is incorporated into the GC approach in our strategy.

In our method, a shape term (a new energy term) is attached to the GC energy function for optimization. To acquire the shape prior, the results of ASM are used. As mentioned above, ASM gives the mean shape and statistical variability. To define the shape term, it is assumed that the largest intensity difference between neighboring voxels is in the direction normal to the object edge so that only the vertical components of the v modes are used. Figure 4 shows the mode projection. For a landmark, the mode is drawn as a vector. After determination of the normal direction of this landmark according to the triangulation information, the vector is projected to the normal direction. This vertical component (yellow arrow) is used to construct the shape term.

2.4.3 Graph-Cut with Shape Prior

The shape term is defined in terms of statistical variability and added to the original GC energy function to improve both data and boundary information. Intuitively, an edge that is closer to the mean shape boundary will be more likely to be cut. The two neighboring vertices should be then distributed to different components. Similarly, a vertex whose Euclidean distance to the desired boundary is large enough belongs to the background or the object depending on whether it is inside the mean surface. Furthermore, a vertex quite close to the mean shape has an unclear label, on which occasion data and boundary term will provide guidance for label assignment.

In our SP-GC approach, the energy function for minimization is defined as:

$$C(L) = \lambda(k_1 D(L) + (1 - k_1) S_1(L)) + (1 - \lambda)(k_2 B(L) + (1 - k_2) S_2(L)) \quad (9)$$

where $S_1(L)$ and $S_2(L)$ are the two shape terms for supplying data and boundary information. k_1 and k_2 are positive parameters. The definition of $S_1(L)$ is similar to that of $D(L)$:

$$S_1(L) = \sum_{p \in V} Pr_p(L_p) \quad (10)$$

First, the distance between vertex p and the mean shape \bar{S} , which is denoted as $\text{Dist}(p, \bar{S})$, is defined as the distance between p and the nearest vertex describing \bar{S} . To define $Pr_p(L_p)$ for $S_1(L)$, with user-defined factor T_n , several situations are considered: When $\text{Dist}(p, \bar{S}) > T_n$, if p is inside \bar{S} :

$$\begin{cases} Pr_p(0) = Inf \\ Pr_p(1) = 0 \end{cases} \quad (11)$$

Otherwise,

$$\begin{cases} Pr_p(0) = 0 \\ Pr_p(1) = Inf \end{cases} \quad (12)$$

When $\text{Dist}(p, \bar{S}) \leq T_n$:

$$\begin{cases} Pr_p(0) = \exp\left(\frac{-\text{Dist}(p, \bar{S})^2}{2\sigma_{S,a}^2}\right) \\ Pr_p(1) = 1 - \exp\left(\frac{-\text{Dist}(p, \bar{S})^2}{2\sigma_{S,a}^2}\right) \end{cases} \quad (13)$$

where a is the vertex index in \bar{S} which is closest to p and Inf is infinity. T_n is a parameter used to distinguish confirmable vertices from those difficult to allocate labels. When a vertex is far away from the mean surface boundary (with distance larger than T_n), its label can be determined in advance. For example, if p is inside \bar{S} and

Fig. 3 Illustration of shape unification. Scanning ranges of datasets in **a** and **b** are different. After definition of individual cutting planes (white line in **a** and **b**), proximal femurs are registered (five surfaces are illustrated in white, red, green, yellow, and blue) and then unified plane is found to give final shape unification result

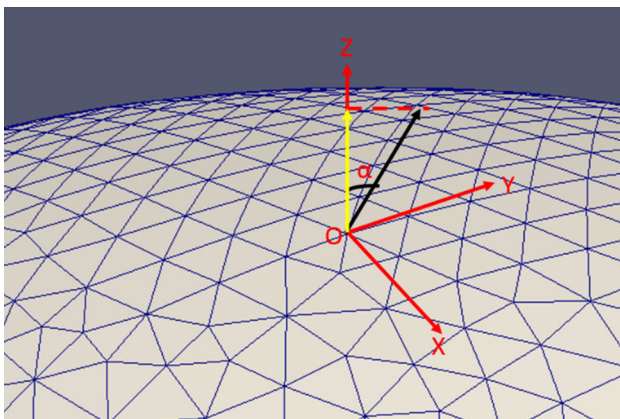
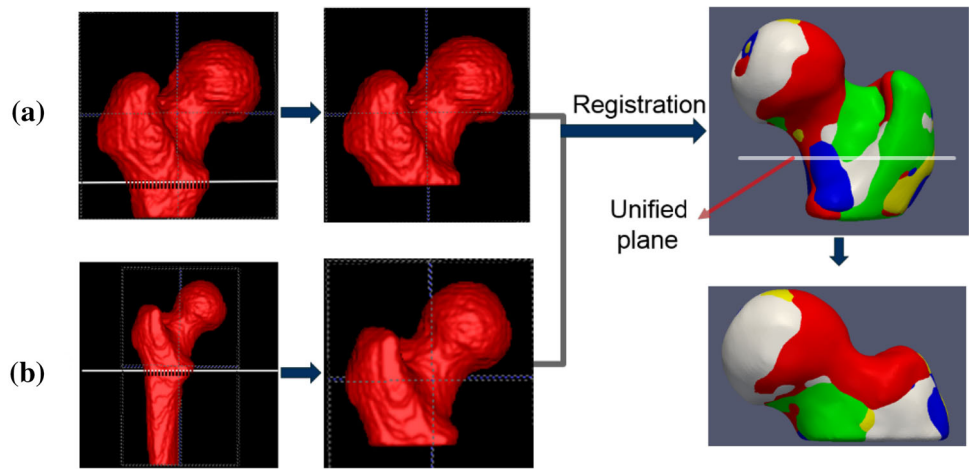


Fig. 4 Mode of one landmark shown as black arrow line. The mode is projected to its normal direction related to mean surface (yellow arrow line)

$Dist(p, \bar{S}) > T_n$, $Pr_p(1)$ will equal 0 and $Pr_p(0) = Inf$, which means that vertex p will be allocated as the object.

The definition of $S_2(L)$ is related to neighboring nodes p_1 and p_2 as:

$$S_2(L) = \sum_{\langle p_1, p_2 \rangle \in E} (\delta(L_{p_1}, L_{p_2}) S_{2_{\langle p_1, p_2 \rangle}}) \tag{14}$$

where $\delta(L_{p_1}, L_{p_2})$ has the same definition as before and:

$$S_{2_{\langle p_1, p_2 \rangle}} \propto \left(1 - \frac{1}{2} \sum_{i=1}^2 \exp\left(\frac{-Dist(p_i, \bar{S})^2}{2\sigma_{S,a}^2}\right) \right) \tag{15}$$

When the distance between p_1 and \bar{S} is very small, as p_2 is adjacent to p_1 (at a distance of 1 voxel in a 4-connected system), $Dist(p_2, \bar{S})$ should be also small. In such circumstances, $S_{2_{\langle p_1, p_2 \rangle}}$ is small enough to approach 0; that is, the edge connecting these two vertices is more likely to be cut and each vertex is assigned different labels.

In the proposed SP-GC strategy, an important parameter is $\sigma_{S,a}$, which is derived from the ASM training stage. From Sect. 2.4.2, for all the landmarks, each mode is projected to the direction perpendicular to the mean shape. Thus, the shape prior can be embedded into the definition of $\sigma_{S,a}$ as follows:

$$\sigma_{S,a} = \frac{1}{k} \sum_{i=1}^v L_{a,i} \tag{16}$$

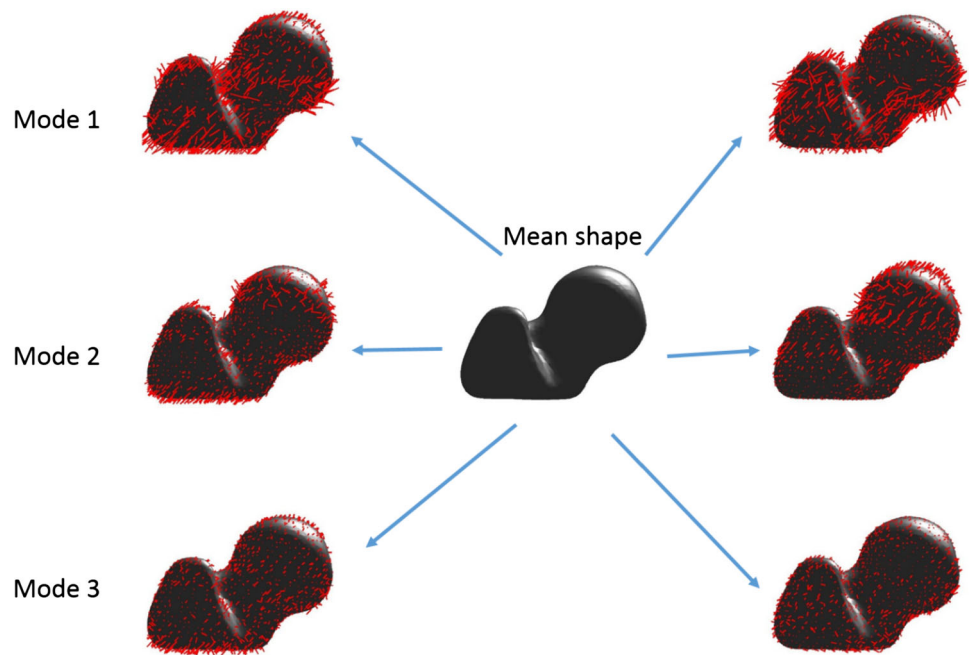
where $a = 1, 2, \dots, n$ and $L_{a,i}$ is the length of the i th mode for landmark index a . k is a positive factor used to determine the proportion of the shape prior used in experiments. The calculation of $\sigma_{S,a}$ generalizes the v statistical modes summarized by ASM construction and is attached to the GC energy function (Fig. 5). The optimization of the objective function is achieved with the min-cut/max-flow algorithm. Although it is only able to find approximate solutions for multi-label GC problems, in our foreground/background image segmentation, the algorithm finds a global optimum for binary labeling problems.

2.5 Experiments

A total of 84 femurs were included in this study. For both training and segmentation stages, all 84 femurs were segmented with ITK-SNAP (<http://www.itksnap.org/>) in a semi-automatic manner. Reliable segmentation results were needed for training. Manually segmented volumes served as the ground truths for quantitatively evaluating segmentation accuracy. The method was implemented in MATLAB R2011a (The MathWorks Inc., MA).

To obtain the distributions of the background and the object for the data term in the GC-based energy function, usually manual seed selection is necessary [34]. In the proposed segmentation strategy, ASM training is essential

Fig. 5 Illustration of active shape modeling results. First three principal modes of statistical variability are represented by *red solid lines* on mean shape. *Mode 1* captures the most prominent statistical variability



for shape term construction. This step facilitates the acquisition of intensity distributions with segmented binary masks of training samples. In our experiments, three segmentation methods, namely ASM, original GC, and SP-GC, were compared with the following measurements: DSC ($DSC = 2|A \cap M| / (|A| + |M|)$), $Diff(A, M) = (|A| - |A \cap M|) / |A|$, and $Diff(M, A) = (|M| - |A \cap M|) / |M|$, where $|A|$ and $|M|$ refer to the volumes of the segmentation results of the proposed method and ground truth, respectively, and $|A \cap M|$ indicates their overlapping volumes. The average surface distance, which refers to the average distance between vertices on the automatically segmented surface and the ground truth surface, was also used. DSC is a commonly used measurement to reflect the overlap ratio of segmentation results and the ground truth. With perfect image segmentation, DSC is 1; with no overlap, it is 0. For the other three measurements, namely $Diff(A, M)$, $Diff(M, A)$, and average surface distance, the smaller their value is, the more accurate is the segmentation result.

In the proposed method, two shape priors corresponding to data and boundary information, respectively, were embedded into the GC. The influence of each term and that of the complete SP-GC were measured using the four measurements.

Both ASM and the SP-GC algorithm need data training to incorporate the shape prior to guide image segmentation. The impact of the number of training samples on segmentation accuracy was evaluated. 12, 16, 20, 24, 28, 32, 36, and 40 training samples were tested for both ASM and the proposed method. The four measurements mentioned above were applied to evaluate the results. The parameters

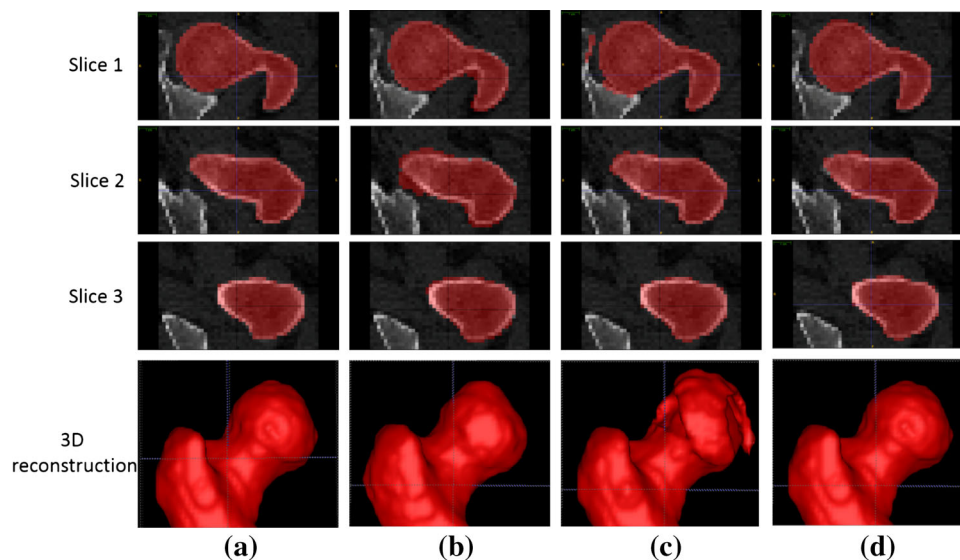
used in the experiments were set as follows: $\lambda = k_1 = k_2 = 0.5$, $\sigma = 30$, $\alpha_v = 95\%$, $T_n = 5$, and $k = 7$.

3 Results

3.1 Segmentation Accuracy

A qualitative comparison of ASM, GC, and SP-GC segmentation results of the proximal femur is shown in Fig. 6, which was obtained with a training group of 12 samples for ASM and SP-GC. The results in the first row demonstrate obvious segmentation error with GC for the gap between the femoral head and the acetabulum. This might occur due to the narrow space between these two bony structures. Parts of the acetabular cortical bone were incorrectly treated as the femur. The cortical bone of the acetabulum has a distribution similar to that of femoral bone. There is an intensity difference between cortical and trabecular bone. Without the assistance of the shape prior, GC was easily trapped in a local optimum. Segmentation error was observed with ASM in the second slice. This outcome might be caused by the slightly deficient training samples. For the slice closer to the femoral shaft, all three approaches presented acceptable results, as intensity distinction was obvious. From the 3D reconstruction results, the segmentation error was visualized clearly. As mentioned, the GC algorithm improperly allocated some voxels belonging to the acetabulum to the femur, especially at the narrow part between these two structures. The output from ASM was acceptable except for the smoothed surface and some

Fig. 6 Comparison of three segmentation methods. **a** Ground truth and results of **b** ASM, **c** GC, and **d** SP-GC. First three rows present illustrations in a planar dimension with segmentation results overlaid on original images and last row is 3D reconstruction of segmented objects



loss of details. The results of the SP-GC algorithm were the most accurate.

To evaluate algorithm performance quantitatively, the measurements defined above were applied to 60 testing data using the manual segmentation results as the ground truths. The means and standard deviations (STD) of DSC, Diff(A,M), Diff(M,A), and average shape distance were computed. They are presented in Table 1. With a training set comprising 12 samples, the proposed SP-GC algorithm had a mean DSC of 0.9600, and those of ASM and GC were 0.8769 and 0.9358, respectively. The mean Diff(A,M) was 0.0412 for SP-GC, and those for ASM and GC were larger than 0.09. Although Diff(M,A) has similar property with Diff(A,M) to reflect segmentation error, it is interesting that Diff(M,A) of SP-GC was 0.0383, which is slightly larger than that of GC (0.0257). These results confirm the qualitative results, where GC tended to produce over-segmentation. The average shape distance for SP-GC was smallest (0.885 mm), which is 142 % and 30 % lower than those of ASM and GC, respectively. SP-GC outperformed the other methods. From Table 1, the stability of SP-GC is highlighted with a mean STD value of 0.0130 for DSC, compared with 0.0539 and 0.0150 for ASM and GC, respectively. GC and SP-GC have higher accuracy than that of ASM. Although the stabilities of GC and SP-GC are similar, SP-GC provided better results

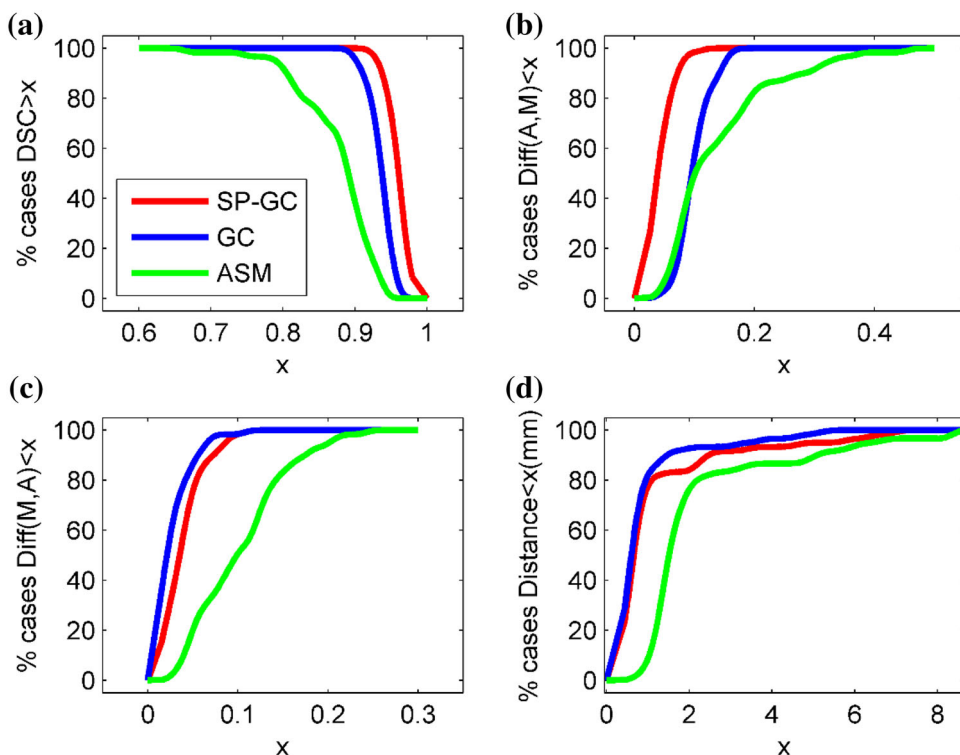
with a larger mean DSC. From the results of statistical significance testing with any two methods related to the performance measurements, there is no statistically significant difference between GC and SP-GC regarding the average shape distance, with a p value of 0.2262, while for the other pairs of methods p -values of <0.05 were obtained.

Another interpretation of segmentation reliability is shown in Fig. 7. The plots respectively summarize the proportions of the 60 testing data with DSC larger than a specified value and Diff(A,M), Diff(M,A), and average shape distance smaller than the value. All 60 cases reached a DSC value of 0.6 for all approaches. As the situation became rigorous, a smaller number of cases satisfied the larger specified DSC value. For instance, when the DSC truncation was 0.95, there were still more than 80 % cases with DSC larger than 0.95 using SP-GC, while for ASM and GC, the ratio decreased rapidly to less than 20 % approximately. It is interesting to note that the plots of Diff(A,M) and Diff(M,A) show a distinguishable tendency which was matched with the results in Table 1 for ASM and SP-GC. When the truncation value on the x-axis was small, such as 0.1, more cases compared to a larger truncation value (e.g. 0.9) using SP-GC had Diff(A,M) smaller than this value, while considering the value of Diff(M,A), an opposite trend could be perceived to support GC as a

Table 1 Quantitative segmentation accuracy obtained using ASM, GC, and SP-GC with 12 training samples

Method	DSC		Diff(A,M)		Diff(M,A)		Distance (mm)	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
ASM	0.8769	0.0539	0.1369	0.0894	0.1015	0.0518	2.148	1.783
GC	0.9358	0.0150	0.0990	0.0260	0.0257	0.0207	1.154	1.438
SP-GC	0.9600	0.0130	0.0412	0.0190	0.0383	0.0221	0.885	0.933

Fig. 7 Segmentation accuracy of *ASM*, *GC*, and *SP-GC*. Plots show overall segmentation performance of 60 testing data with 12 training samples. **a** DSC, **b** Diff(A,M), **c** Diff(M,A), and **d** average distance



better method. The curve of the average shape distance is similar to that of Diff(M,A).

3.2 Segmentation Accuracy with Various Shape Priors

The different segmentation accuracy using SP-GC with each shape prior, namely $S_1(L)$ and $S_2(L)$, based on 12 training samples were given in Table 2. With only incorporate shape prior $S_1(L)$ ($k_2 = 1$) or $S_2(L)$ ($k_1 = 1$) into GC, the segmentation performance was improved, which was however poorer than using the complete SP-GC with both $S_1(L)$ and $S_2(L)$ ($k_1 \neq 1$ and $k_2 \neq 1$). From a numerical view, comparing the impact of $S_1(L)$ and $S_2(L)$, the influence is similar with about 0.02 DSC improvements than original GC. With statistical significance testing, regarding DSC, Diff(A,M), Diff(M,A) and average shape distance, the differences with only $S_2(L)$ and $S_2(L)$ are not significant with p-values of 0.7291, 0.7071, 0.9286 and 0.9797, reflecting the similar effects of these two terms

independently. Refer to the differences of $S_2(L)$ -GC vs. SP-GC, the p-values are 0.007, 0.0027, 0.6733 and 0.4869; $S_1(L)$ -GC vs. SP-GC, the p-values are 0.0039, 0.0007, 0.6725 and 0.5023.

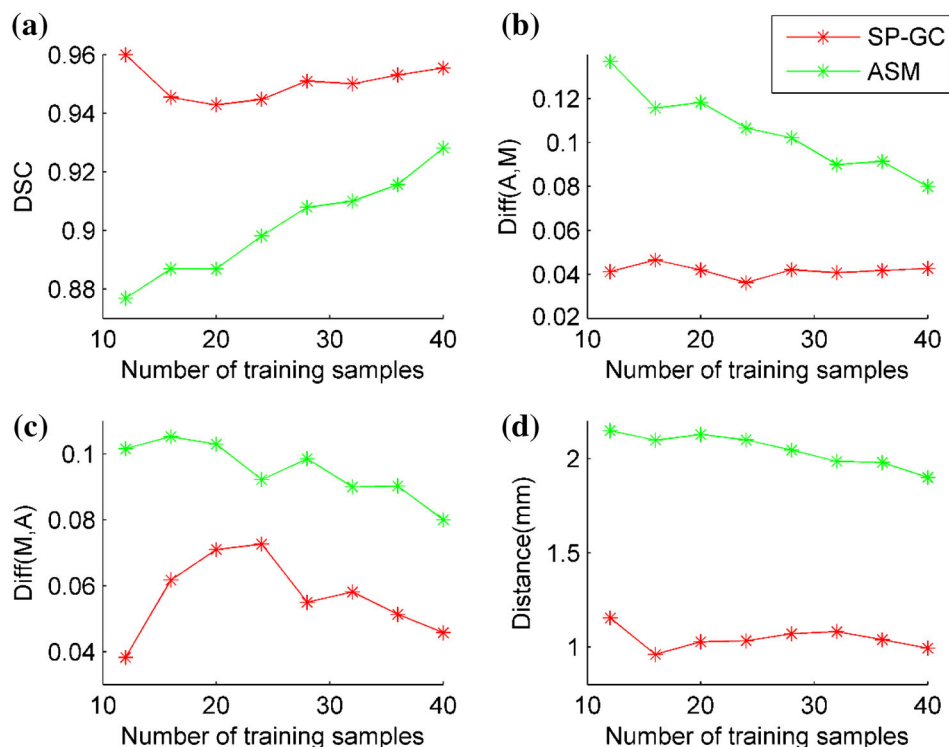
3.3 Effect of Number of Training Samples

To determine the effect of the number of training samples on segmentation accuracy with DSC, Diff(A,M), Diff(M,A), and average shape distance, experiments were carried out with training set sizes of 12, 16, 20, 24, 28, 32, 36, and 40. All other parameters were kept constant. The results are shown in Fig. 8. The values of vertical axis were the means of corresponding measurements. Intuitively, SP-GC was more accurate than ASM even with an increasing sample size according to the DSC, Diff(A,M), Diff(M,A) and average shape distance. For the ASM algorithm, the segmentation accuracy increased with more data used in the training stage (DSC increased and Diff(A,M), Diff(M,A), and average shape distance decreased). DSC

Table 2 Segmentation accuracy obtained using SP-GC with each shape prior based on 12 training samples

Method	DSC		Diff(A,M)		Diff(M,A)		Distance (mm)	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
With S_1	0.9535	0.0131	0.0524	0.0212	0.0400	0.0214	0.977	1.337
With S_2	0.9526	0.0145	0.0539	0.0209	0.0401	0.0255	0.983	1.336
S_1 and S_2	0.9600	0.0130	0.0412	0.0190	0.0383	0.0221	0.885	0.933

Fig. 8 Relationships between number of training samples and **a** DSC, **b** Diff(A,M), **c** Diff(M,A), and **d** average distance. Accuracy of SP-GC is higher than that of ASM



decreased when the training size increased from 12 to 20 for SP-GC, but then became stable. Although Diff(M,A) fluctuated with number of training samples, Diff(A,M) and average shape distance became stable. The number of training samples significantly influences the segmentation accuracy of ASM but not that of the proposed SP-GC. Even with a small number of training samples (e.g., 12), SP-GC had reliable segmentation results.

4 Discussion

This paper presented a scheme for extracting the proximal femur, including femoral head and head-and-neck junction, which is of great significance for diagnosing FAI, based on CT scans. In the proposed framework, the shape prior is obtained using ASM and incorporated into the GC algorithm. The energy function of SP-GC contains not only data and boundary information of the images but also shape information derived from ASM. Accurate segmentation results are achieved by minimizing the energy function with the max-flow/min-cut algorithm.

3D medical image segmentation is increasingly important for extracting information regarding anatomy and function. Image modalities such as MR imaging and CT make it possible for medical personnel to have a more intuitional understanding about a variety of tissues and organs with high resolution for possible 3D reconstruction

[39, 40]. Segmentation is required for further analysis. Traditional manual segmentation is labor- and time-consuming and the results depend on the operator. Many automatic segmentation methods have been proposed. Kang et al. [41] used local adaptive thresholding values to segment femurs in order to avoid incorrectly connecting the femur and the acetabulum using CT data. However, manual intervention is still needed in some situations (the sphere is positioned around the femoral head). Cootes and Taylor proposed using the shape prior as guidance in segmentation. The most typical methods are ASM and the active appearance model. To apply the GC algorithm in image segmentation, the most apparent disadvantage is seed selection by the user [34]. In order to obtain the approximate intensity distribution of the object and the background, users are required to mark some voxels or regions for distribution estimation. In SP-GC, since the intensity can be acquired in the training stage using segmentation results as masks, the intensity distributions of the object and the background can be obtained and more importantly, the estimation results are more reliable. Since automatic segmentation is difficult, some interactive GC methods that use a bounding box to select regions of interest have been proposed for more reliable results [42, 43]. However, merely a bounding box should be too awkward to capture the complete and accurate shape prior in many different applications as the shapes of different structures are not the same. Similar to our work, some

studies have combined shape priors with GC. Slabaugh and Unal [44] used an ellipse to fit a pre-segmented object using conventional GC and extracted the area around the fitted ellipse. For segmentation, an inner band boundary as well as an outer band boundary are defined near the object edge to constrain the solution of GC. However, the weights inside each band are constant, which is unreasonable, based on the assumption in our method that the voxels further away from the boundary should be less possible boundary voxels. In the proposed SP-GC, the weights are determined by a Gaussian function related to the shape information from ASM. In the study of Chen et al. [30], the shape prior was generated from ASM, as is done in our algorithm. However, in their work, the shape prior term is different and is merely embedded into the data term. From the experimental results (Table 2), it can be perceived that the existence of $S_1(L)$ made some contributions to the better performance as X. Chen et al. presented. However, the complete SP-GC, with both $S_1(L)$ and $S_2(L)$, achieved better results than either with merely $S_1(L)$ or $S_2(L)$. From the hypothesis testing, the accuracy differences presented by DSC, Diff(A,M), Diff(M,A) and average shape distance between GC incorporating with $S_1(L)$ and $S_2(L)$ are not statistically significant. Another significant distinction is that their study only provided detailed experimental results for 2D image segmentation. Although the authors stated that their method can be extended to 3D segmentation, no results were given. The shape priors can be embedded into the energy function to reflect data or boundary information, whereas in our scheme, both are considered.

In this study, the shapes for training are unified using a two-step strategy. First, the approximate proximal femur partitions are generated manually, after which rigid registration is performed and a unified plane is acquired to define the final results. The shape prior is calculated as a portion of the sum of vertical components of statistical modes from ASM for each landmark and this information serves as additional information to data and boundary term ($S_1(L)$ and $S_2(L)$). Qualitatively, for 3D reconstruction, GC gave an obvious error for the superior femoral head; it wrongly included voxels of the acetabulum. In spite of a smaller DSC for ASM comparing with GC, the model-based method presented a better 3D visualization. The reason for this might be that during ASM optimization, the updated shape is obtained by transforming the mean shape with variation restriction, which made the deformation of the altered surface less excessive. Although the overlap ratio using ASM was not superior to that of GC, the fundamental anatomy of the femur was preserved. The segmentation accuracy was measured with DSC, Diff(A,M), Diff(M,A), and average shape distance, which showed that the proposed SP-GC method outperformed other methods. For example, the DSC value was 0.9600 for SP-GC, which

is close to the ideal value of 1. Although the DSC value of GC was also outstanding (0.9358), the Diff(A,M) value of GC was more than double that of SP-GC. The poor performance of ASM might be associated with the small number of training samples. With a larger sample size for the training stage, the performance of ASM greatly improved. When more data were included in the training subset, ASM was able to capture more complete shape variability and therefore produced better segmentation results. When the training subset was 12 for SP-GC, the DSC value was 0.9600. Original GC can give a qualitatively acceptable result. With shape prior of segmenting object, the information is more but merely a complementary component for SP-GC and data and boundary terms still play an important role in segmentation. Even with less training samples, it is enough for SP-GC to provide these shape information to guide segmentation with inherent data and boundary information. This finding is important as it reflects the robustness of SP-GC compared to ASM with a low number of training data.

Although the segmentation framework presented in this paper was used for proximal femur segmentation in our experiments, it can be applied for segmenting other structures. With the training stage, seed selection is unnecessary for obtaining the intensity distributions of the object and the boundary for data term definition. With the shape term attached to the GC algorithm, initialization of the mean shape on segmentation data is required. Similar to ASM, the initialization should be approximately close to the real boundary of the object. In our work the initialization was achieved by manually setting the translational, rotational, and scaling parameters heuristically, which is a limitation of the method. With the data and boundary term, SP-GC does not depend on the shape prior completely, and thus the segmentation results are not influenced too much by initializing position. The effect of initialization and a possible automatic initialization method require future investigation.

5 Conclusion

A method for segmenting the proximal femur was proposed to support evidence-based diagnosis and treatment decision. First, the shape prior is obtained from ASM with a mean shape and a number of statistical modes describing shape variability. This information is embedded into the GC approach as a supplementary term in addition to the original data and boundary term. The max-flow/min-cut method is applied to minimize the objective function. Compared with ASM and GC, the proposed SP-GC approach achieves better segmentation accuracy in terms of a larger DSC and smaller Diff(A,M), Diff(M,A), and

average shape distance. SP-GC is less sensitive to the setting of the number of training samples in comparison with traditional ASM.

Acknowledgments This work was partially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No.: CUHK 473012).

References

- Ganz, R., Parvizi, J., Beck, M., Leunig, M., Nötzli, H., & Siebenrock, K. A. (2003). Femoroacetabular impingement: a cause for osteoarthritis of the hip. *Clinical Orthopaedics and Related Research*, 417, 112–120.
- Ganz, R., Leunig, M., Leunig-Ganz, K., & Harris, W. H. (2008). The etiology of osteoarthritis of the hip. *Clinical Orthopaedics and Related Research*, 466, 264–272.
- Beaulé, P. E., Zaragoza, E., Motamedi, K., Copelan, N., & Dorey, F. J. (2005). Three-dimensional computed tomography of the hip in the assessment of femoroacetabular impingement. *Journal of Orthopaedic Research*, 23, 1286–1292.
- Clohisey, J. C., Carlisle, J. C., Beaulé, P. E., Kim, Y. J., Trousdale, R. T., Sierra, R. J., et al. (2008). A systematic approach to the plain radiographic evaluation of the young adult hip. *Journal of Bone Joint Surgery*, 90, 47–66.
- Rakhra, K. S., Sheikh, A. M., Allen, D., & Beaulé, P. E. (2009). Comparison of MRI alpha angle measurement planes in femoroacetabular impingement. *Clinical Orthopaedics and Related Research*, 467, 660–665.
- Harris, M. D., Datar, M., Whitaker, R. T., Jurrus, E. R., Peters, C. L., & Anderson, A. E. (2013). Statistical shape modeling of cam femoroacetabular impingement. *Journal of Orthopaedic Research*, 31, 1620–1626.
- Raut, S., Raghuvanshi, M., Dharaskar, R., & Raut, A. (2009). Image segmentation—a state-of-art survey for prediction. *Proceedings IEEE International Conference Advanced Computer Control* (pp. 420–424).
- Naz, S., Majeed, H., & Irshad, H. (2010). Image segmentation using fuzzy clustering: A survey. *Proceedings IEEE International Conference Emerging Technologies*, (pp. 181–186).
- Du, M., Ding, Y., & Jia, Q. (2013). A multi-threshold segmentation method based on ant colony algorithm. *5th International Conference Machine Vision* (pp. 878402–878409).
- Xiao, B., Jing, Y., & Guan, Y. (2013). A novel automatic thresholding segmentation method with local adaptive thresholds. *arXiv preprint*, arXiv: 1305.5160.
- Yi, F., & Moon, I. (2012). Image segmentation: A survey of graph-cut methods. *Proceedings IEEE International Conference Systems and Informatics* (pp. 1936–1941).
- Felzenszwalb, P. F., & Huttenlocher, D. P. (2004). Efficient graph-based image segmentation. *Journal of Computer Vision*, 59, 167–181.
- Sundaramoorthi, G., Yezzi, A., & Mennucci, A. C. (2008). Coarse-to-fine segmentation and tracking using sobolev active contours. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30, 851–864.
- Willett, R. M., & Nowak, R. D. (2007). Minimax optimal level-set estimation. *IEEE Transactions on Image Processing*, 16, 2965–2979.
- Falcão, A. X., Udupa, J. K., & Miyazawa, F. K. (2000). An ultra-fast user-steered image segmentation paradigm: Live wire on the fly. *IEEE Transactions on Medical Imaging*, 19, 55–62.
- Cootes, T. F., Taylor, C. J., Cooper, D. H., & Graham, J. (1995). Active shape models—their training and application. *Computer Vision and Image Understanding*, 61, 38–59.
- Cootes, T. F., Edwards, G. J., & Taylor, C. J. (2001). Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23, 681–685.
- Yokota, F., Okada, T., Takao, M., Sugano, N., Tada, Y., & Tomiyama, N. et al. (2013). Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models. *International Conference Medical Image Computing and Computer-Assisted Intervention* (pp. 190–197).
- Wang, D., Shi, L., Chu, W. C., Cheng, J. C., & Heng, P. A. (2009). Segmentation of human skull in MRI using statistical shape information from CT data. *Journal of Magnetic Resonance Imaging*, 30, 490–498.
- Boykov, Y. Y., & Kolmogorov, V. (2004). An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26, 1124–1137.
- Krcch, M., Székely, G., & Blanc, R. (2011). Fully automatic and fast segmentation of the femur bone from 3D-CT images with no shape prior. *Proceedings IEEE International Conference Biomedical Imaging: From Nano to Macro* (pp. 2087–2090).
- Freedman, D., & Zhang, T. (2005). Interactive graph cut based segmentation with shape priors. *Proceedings IEEE International Conference Computer Vision and Pattern Recognition*, 1, 755–762.
- Zoroofi, R. A., Sato, Y., Sasama, T., Nishii, T., Sugano, N., Yonenobu, K., et al. (2003). Automated segmentation of acetabulum and femoral head from 3-D CT images. *IEEE Transactions on Information Technology in Biomedicine*, 7, 329–343.
- Cheng, Y., Zhou, S., Wang, Y., Guo, C., Bai, J., & Tamura, S. (2013). Automatic segmentation technique for acetabulum and femoral head in CT images. *Pattern Recognition*, 46, 2969–2984.
- O’Neill, G. T., Lee, W. S., & Beaulé, P. (2012). Segmentation of cam-type femurs from CT scans. *The Vision Computing*, 28, 205–218.
- Nakagomi, K., Shimizu, A., Kobatake, H., Yakami, M., Fujimoto, K., & Togashi, K. (2013). Multi-shape graph cuts with neighbor prior constraints and its application to lung segmentation from a chest CT volume. *Medical Image Analysis*, 17, 62–77.
- Essa, E., Xie, X., Sazonov, I., Nithiarasu, P., & Smith, D. (2013). Shape prior model for media-adventitia border segmentation in IVUS using graph cut. *Medical Computer Vision*, 7766, 114–123.
- Xiong, W., Li, A. L., Ong, S. H., & Sun, Y. (2013). Automatic 3D prostate MR image segmentation using graph cuts and level sets with shape prior. *International Conference Advances in Multimedia Information Processing* (pp. 211–220).
- Wang, H., Zhang, H., & Ray, N. (2013). Adaptive shape prior in graph cut image segmentation. *Pattern Recognition*, 46, 1409–1414.
- Chen, X., Udupa, J. K., Alavi, A., & Torigian, D. A. (2013). GC-ASM: Synergistic integration of graph-cut and active shape model strategies for medical image segmentation. *Computer Vision and Image Understanding*, 117, 513–524.
- Belongie, S., Malik, J., & Puzicha, J. (2002). Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24, 509–522.
- Heinonen, T., Eskola, H., Dastidar, P., Laarne, P., & Malmivuo, J. (1997). Segmentation of T1 MR scans for reconstruction of resistive head models. *Computer Methods and Programs in Biomedicine*, 54, 173–181.
- Gonzalez, G. C., & Woods, R. E. (1992). *Digital imaging processing*. Cambridge, MA: Wesley.

34. Boykov, Y. Y., & Jolly, M. P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in ND images. *Proceedings IEEE International Conference Computer Vision*, 1, 105–112.
35. Lorensen, W. E., & Cline, H. E. (1987). Marching cubes: A high resolution 3D surface construction algorithm. *International Conference ACM Siggraph Computer Graphics*, 21, 163–169.
36. van Ginneken, B., Frangi, A. F., Staal, J. J., ter Haar Romeny, B. M., & Viergever, M. A. (2002). Active shape model segmentation with optimal features. *IEEE Transactions on Medical Imaging*, 21, 924–933.
37. Golub, G. H., & Reinsch, C. (1970). Singular value decomposition and least squares solutions. *Numerische Mathematik*, 14, 403–420.
38. Besl, P. J., & McKay, N. D. (1992). Method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, 239–256.
39. Diederichs, G., Seim, H., Meyer, H., Issever, A. S., Link, T. M., Schröder, R. J., & Scheibel, M. (2008). CT-based patient-specific modeling of glenoid rim defects: a feasibility study. *American Journal of Roentgenology*, 191, 1406–1411.
40. Shi, L., Wang, D., Chu, W. C., Burwell, G. R., Wong, T. T., Heng, P. A., & Cheng, J. C. (2011). Automatic MRI segmentation and morphoanatomy analysis of the vestibular system in adolescent idiopathic scoliosis. *Neuroimage*, 54, S180–S188.
41. Kang, Y., Engelke, K., & Kalender, W. A. (2003). A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. *IEEE Transactions on Medical Imaging*, 22, 586–598.
42. Lempitsky, V., Kohli, P., Rother, C., & Sharp, T. (2009). Image segmentation with a bounding box prior. *Proceedings IEEE International Conference Computer Vision* (pp. 277–284).
43. Peng, Y., & Liu, R. (2010). Object segmentation based on watershed and graph cut. *Proceedings IEEE International Conference Image and Signal Processing*, 3, 1431–1435.
44. Slabaugh, G., & Unal, G. (2005). Graph cuts segmentation using an elliptical shape prior. *Proceedings IEEE International Conference Image Processing*, 2, 1222–1225.