

# Statistical model-based segmentation of the proximal femur in digital antero-posterior (AP) pelvic radiographs

Weiguo Xie · Jochen Franke · Cheng Chen ·  
Paul A. Grützner · Steffen Schumann ·  
Lutz-P. Nolte · Guoyan Zheng

Received: 11 January 2013 / Accepted: 16 July 2013 / Published online: 31 July 2013  
© CARS 2013

## Abstract

**Purpose** Segmentation of the proximal femur in digital antero-posterior (AP) pelvic radiographs is required to create a three-dimensional model of the hip joint for use in planning and treatment. However, manually extracting the femoral contour is tedious and prone to subjective bias, while automatic segmentation must accommodate poor image quality, anatomical structure overlap, and femur deformity. A new method was developed for femur segmentation in AP pelvic radiographs.

**Methods** Using manual annotations on 100 AP pelvic radiographs, a statistical shape model (SSM) and a statistical appearance model (SAM) of the femur contour were constructed. The SSM and SAM were used to segment new AP pelvic radiographs with a three-stage approach. At initialization, the mean SSM model is coarsely registered to the femur in the AP radiograph through a scaled rigid registration. Mahalanobis distance defined on the SAM is employed as the search criteria for each annotated suggested landmark

location. Dynamic programming was used to eliminate ambiguities. After all landmarks are assigned, a regularized non-rigid registration method deforms the current mean shape of SSM to produce a new segmentation of proximal femur. The second and third stages are iteratively executed to convergence.

**Results** A set of 100 clinical AP pelvic radiographs (not used for training) were evaluated. The mean segmentation error was  $0.96 \text{ mm} \pm 0.35 \text{ mm}$ , requiring  $< 5 \text{ s}$  per case when implemented with Matlab. The influence of the initialization on segmentation results was tested by six clinicians, demonstrating no significance difference.

**Conclusions** A fast, robust and accurate method for femur segmentation in digital AP pelvic radiographs was developed by combining SSM and SAM with dynamic programming. This method can be extended to segmentation of other bony structures such as the pelvis.

**Electronic supplementary material** The online version of this article (doi:10.1007/s11548-013-0932-5) contains supplementary material, which is available to authorized users.

W. Xie (✉) · C. Chen · S. Schumann · L.-P. Nolte · G. Zheng  
Institute for Surgical Technology and Biomechanics,  
University of Bern, Stauffacherstrasse 78, 3014 Bern, Switzerland  
e-mail: weiguo.xie@istb.unibe.ch

G. Zheng  
e-mail: guoyan.zheng@ieee.org

W. Xie  
Graduate School for Cellular and Biomedical Sciences,  
University of Bern, Bern, Switzerland

W. Xie · J. Franke · P. A. Grützner  
BG Trauma Centre Ludwigshafen at Heidelberg University Hospital,  
Ludwigshafen, Germany

**Keywords** Statistical shape model · Statistical appearance model · Segmentation · AP pelvic radiograph · Dynamic programming

## Abbreviations

AP	Antero-posterior
3D	Three-dimensional
2D	Two-dimensional
SSM	Statistical shape model
SAM	Statistical appearance model
THA	Total hip arthroplasty
DHS	Dynamic hip screw
PFN	Proximal femur nail
CT	Computed tomography
MRI	Magnetic resonance imaging

OA	Osteoarthritis
FAI	Femoroacetabular impingements
TVO	Transtrochanteric valgus osteotomy
PCA	Principal component analysis

## Introduction

Digital radiography is a widely used 2D imaging modality for disease diagnosis, medical treatment and surgical planning. Recently Zheng [1,2] introduced a method to reconstruct a patient-specific three-dimensional (3D) surface model of the pelvis from a single two-dimensional (2D) antero-posterior (AP) pelvic radiograph. There are several clinical challenges, where a patient-specific 3D model of the hip joint might be helpful. Such a 3D model can be used by surgeons to understand and simulate patient's joint mechanics such as range of motion [3] or to plan an intervention such as a total hip arthroplasty (THA). A 3D model of the proximal femur may also be used for computer-assisted navigation in orthopaedic trauma surgery to implant, for instance, a dynamic hip screw (DHS) (on-line supplemental Fig. 1a) or a proximal femur nail (PFN) (on-line supplemental Fig. 1b). Such a patient-specific model like those generated from other imaging modalities such as computed tomography (CT) and magnetic resonance imaging (MRI) also benefits the post-operative evaluation of implanted stem and head of THA [3] (on-line supplemental Fig. 1c), and the quantitative analysis of hip osteoarthritis (OA) [3–6], including the morphological assessment of femoroacetabular impingements (FAI) [5,6]. In order to achieve the aforementioned benefits from the 3D model built with the methods introduced by Zheng [1,2], an accurate segmentation of the proximal femur in 2D AP pelvic X-ray radiographs is the first and also a crucial step [1,2].

It is possible to manually extract contours of the proximal femur in AP pelvic X-ray radiographs. This process, however, is time-consuming, tedious and prone to subjective bias. Thus, an automated solution will facilitate the clinical integration of the methods introduced by Zheng [1,2]. Development of an automated solution for segmenting the proximal femur in AP pelvic X-ray radiographs is challenging due to variation in image quality, overlap of neighbouring structures and anatomical deformities.

Due to the pervasive applications of AP pelvic radiographs in orthopaedic diagnosis and surgical planning, attempts to develop an automatic solution for proximal femur segmentation have been reported [7–13]. Behiels et al. [7] and Pilgram et al. [8] utilized statistical appearance models (SAM) and statistical shape models (SSM) or their extensions, and applied both models in the scenario of femur segmentation. In Lindner et al. [9], random forest regression voting was employed to achieve an accuracy of approximately 1 mm that was claimed as the most accurate automatic method for seg-

menting femur in AP pelvic radiographs yet reported [9]. Chen et al. [10] proposed to initialize a 2D model by searching parallel lines and circles in regions of femoral shaft and head, respectively. Then, this 2D model was deformed with a non-rigid registration algorithm. In Boukala et al. [11], the mean shape of SSM was registered to the input image by explicitly finding correspondence with shape context descriptor. The registered mean shape was then deformed to refine the segmented femoral contour. Recently, Gamage et al. [12] described a method for femur segmentation, which non-rigidly registered the projected silhouette of a generic 3D femur model with the elaborately extracted edges of radiographs to achieve the segmentation. In case the radiograph was acquired in an abnormal acquisition pose, this method required a manual pose initialization of the generic 3D model regarding position, orientation and scale. Furthermore, Ding et al. [13] presented an atlas-based approach consisting of two stages: global alignment and local refinement. The bone models in the atlas were globally aligned with the detected edges in the image. After global alignment, a level set method with shape constraint was applied to deform the femur model to extract detailed femoral contours.

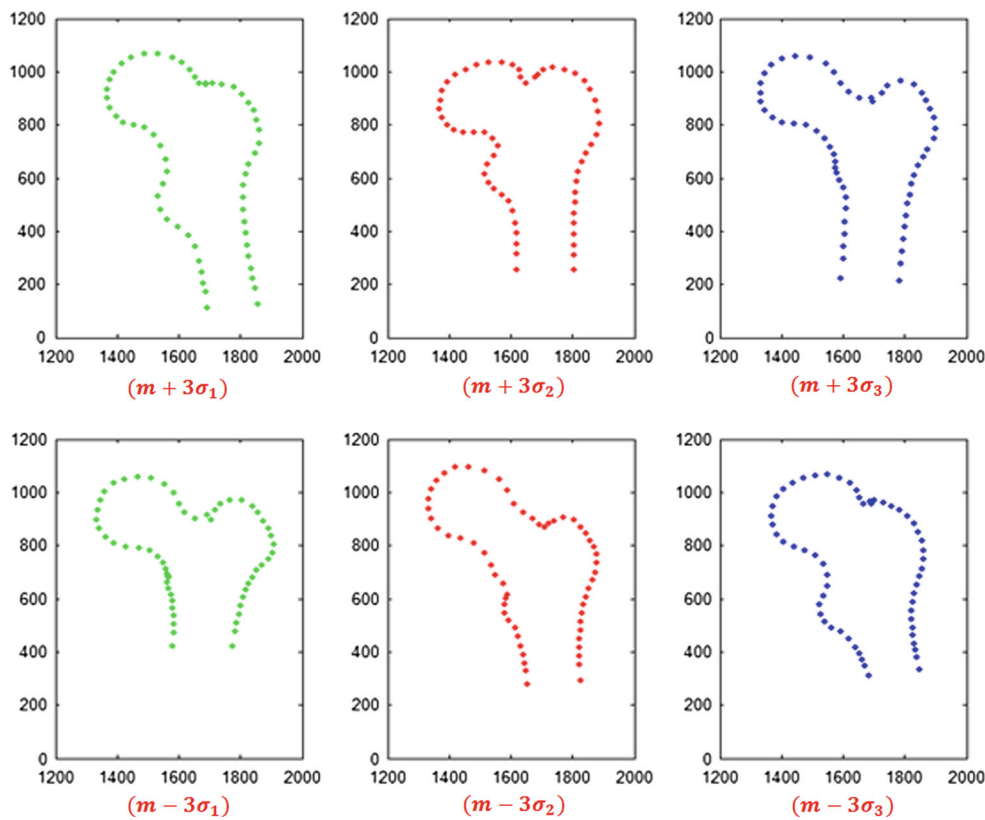
In this paper, a fast, robust and accurate method for segmenting the proximal femur in AP pelvic X-ray radiographs is presented, based on statistical models of shape and grey-level appearance [14,15], in combination with the dynamic programming technique. A comparison of the proposed method with previous work is presented.

## Methods

Prior to describing the details of the segmentation method, we would like to first present the construction of the SSM and of the SAM [14,15] that will be used in our method. In this article, only the method for segmenting the left proximal femur is presented. The same procedure can be applied to the right side.

### Construction of the statistical shape model

One hundred AP pelvic X-ray radiographs with significant variations of femur shape and appearance acquired in routine clinical practice were used to construct the SSM in this paper. To build this model, the left proximal femur was annotated with fifty-nine landmarks by propagating a manual annotation on a reference femur contour to all other contours via a non-rigid registration (please refer to Fig. 2 for the annotated examples). Principal component analysis (PCA) was then applied to obtain the SSM of the left proximal femur. Figure 1 illustrates the variations of three times standard deviation along its three largest eigenmodes (principal components) (all axes in *pixel*).



**Fig. 1** Variations of the three times standard deviation of the constructed proximal femur SSM along the directions of its three largest principal components (all axes in *pixel*)

Construction of the statistical appearance model

Due to the existence of metal objects such as prostheses implanted for hip degeneration or trauma, and fixation screws/nails for femur fracture, a certain number of training images used in the SSM construction had to be excluded for the purpose of constructing the SAM. Consequently, sixty-seven training AP pelvic radiographs with the aligned annotations were used to generate the SAM of the left proximal femur.

All fifty-nine landmarks in the aligned annotations were numbered and located at certain anatomical structures of left femur. For example (Fig. 2), landmarks 21, 29, 37, 42 and 49 were marked at the greater trochanter, upper boundary contour point between femoral neck and head, fovea of femoral head, lower boundary point between neck and head, as well as the lesser trochanter, respectively.

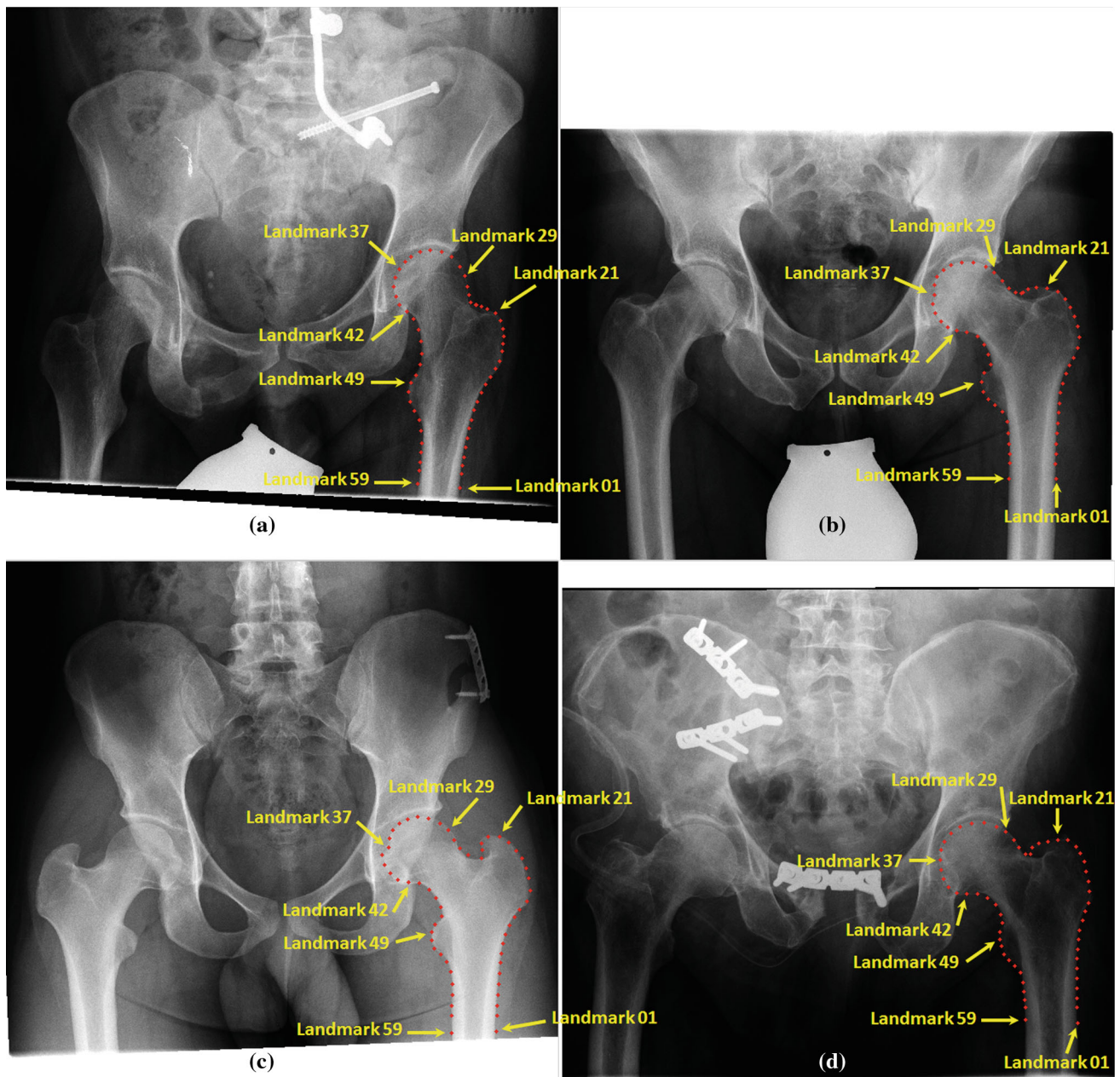
For a landmark (depicted as a red dot in Fig. 3b) in a training image, there exists a normal direction  $\vec{n}$  as shown in Fig. 3b. Along this normal direction, we sampled  $n_t$  points (note: we empirically chose  $n_t$  as twenty). Similarly, we also sampled  $n_t$  points in the inverse direction (the yellow one in Fig. 3b) of  $\vec{n}$ . Thus, in total  $(2n_t + 1)$  points were sampled. On both sides of each  $(2n_t + 1)$  points, further five (note: five

is an empirical value) points were digitized along respective orientations of  $\vec{n}_l$  and  $\vec{n}_r$  (green and cyan in Fig. 3b) that were perpendicular to  $\vec{n}$ . Such a process led to an  $(2n_t + 1) \times 10$  array containing intensities at  $(2n_t + 1) \times 10$  sampled points in the sampling area of one landmark. Figure 3a shows the sampling areas of all landmarks in a training radiograph, and Fig. 3c presents the resulting image when the intensities of all sampled areas are combined into one image.

By averaging intensities of ten points sampled on both sides of each  $(2n_t + 1)$  points, intensities at those  $(2n_t + 1)$  points of each landmark were represented by a  $(2n_t + 1) \times 1$  vector derived from the  $(2n_t + 1) \times 10$  array. Without loss of generality, given a landmark  $i$  in a training radiograph  $j$ , its training appearance model  $g'_{ij}$  was computed as derivatives of intensity values in its  $(2n_t + 1) \times 1$  vector [14, 15]. Dimensions of  $g'_{ij}$  were also  $(2n_t + 1) \times 1$ , and each element of  $g'_{ij}$  needed to be normalized with

$$g_{ij} = \frac{g'_{ij}}{\sum_{k=1}^{2n_t+1} |g'_{ijk}|} \tag{1}$$

For the  $i$ th landmark across the sixty-seven training images, its mean normalized appearance model  $g_i$  was calculated by



**Fig. 2** Examples of the landmark-based annotation of the left proximal femur in the training radiographs

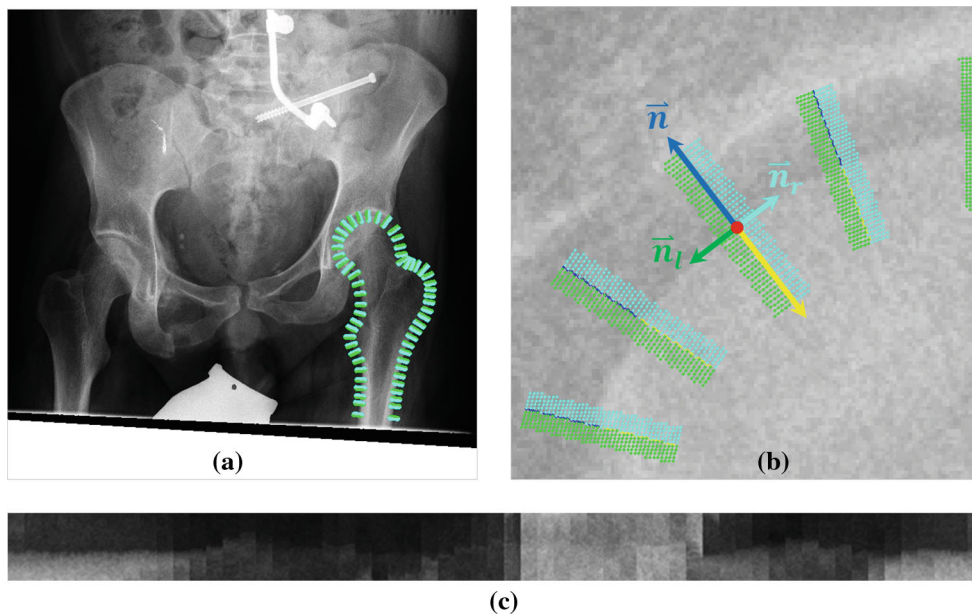
$$g_i = \frac{1}{67} \sum_{j=1}^{67} g_{ij} \quad (2)$$

Based on  $g_i$ , the covariance matrix  $S_{g_i}$  of variations between the training appearance models  $g_{ij}$  and the mean model  $g_i$  could be generated. Then, SAM of each landmark was obtained with PCA. Figure 4 visualizes the mean appearance model  $g_i$  as well as the variations of its three times standard deviation of the SAM along the directions of its three largest eigenmodes for all fifty-nine landmarks.

### Statistical model-based proximal femur segmentation

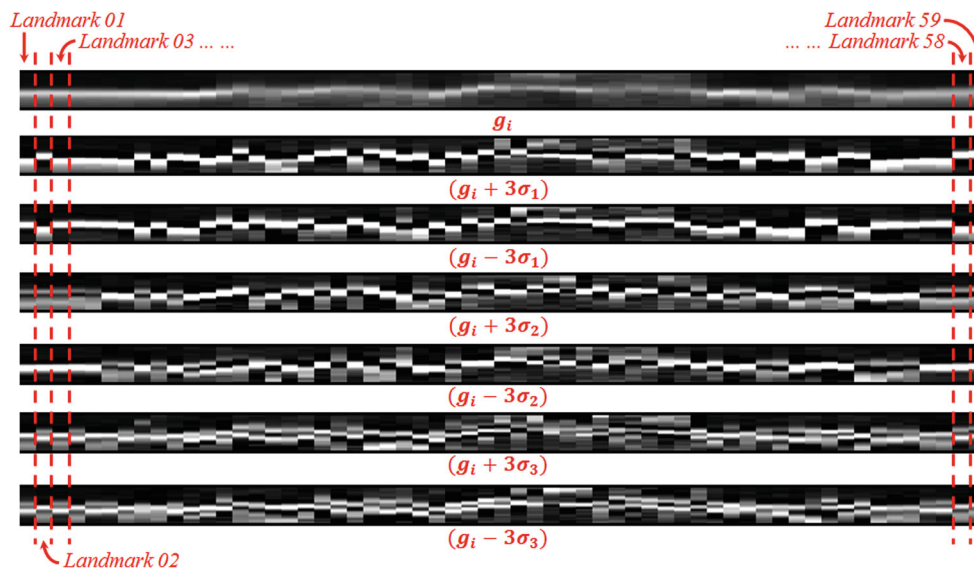
Based on the SSM and the SAM constructed in section “construction of the statistical shape model” and “construction of the statistical appearance model”, our statistical model-based proximal femur segmentation method consists of the following three stages.

1. Initialization stage: register the mean model of the SSM coarsely with the proximal femur in a given test image, using a scaled rigid registration algorithm [16, 17].



**Fig. 3** **a** Sampling areas of all landmarks in a training image. **b** The sampling scheme for each landmark: the normal direction (blue), the inverse direction of the normal (yellow), left side (green) and right side

(cyan) of the normal. **c** Combining the intensities of the sampling areas in (a) into one image



**Fig. 4** The mean appearance model for all fifty-nine landmarks and variations of the three times standard deviation of the constructed SAM along its three largest principal components

2. SAM-based registration stage

- Step 2-1: for each landmark  $i$  on the mean model of the SSM, set a searching range  $n_s$ .
- Step 2-2: for each landmark  $i$  along its normal direction (both positive and negative directions), search  $(2n_s + 1)$  points.
- Step 2-3: within the searched  $(2n_s + 1)$  points, select those points whose intensity derivatives are larger than the half of the maximal derivative as the candidate points for landmark  $i$ .
- Step 2-4: compute the grey-level appearance model of each candidate point, using the same scheme as we introduced in section “construction of the statistical appearance model”.
- Step 2-5: compute Mahalanobis distances [14, 16] of all candidate points and select the candidate point with the minimal Mahalanobis distance as the suggested point for landmark  $i$ .
- Step 2-6: for each landmark that has been annotated as the contour on the femoral head, employ the dynamic

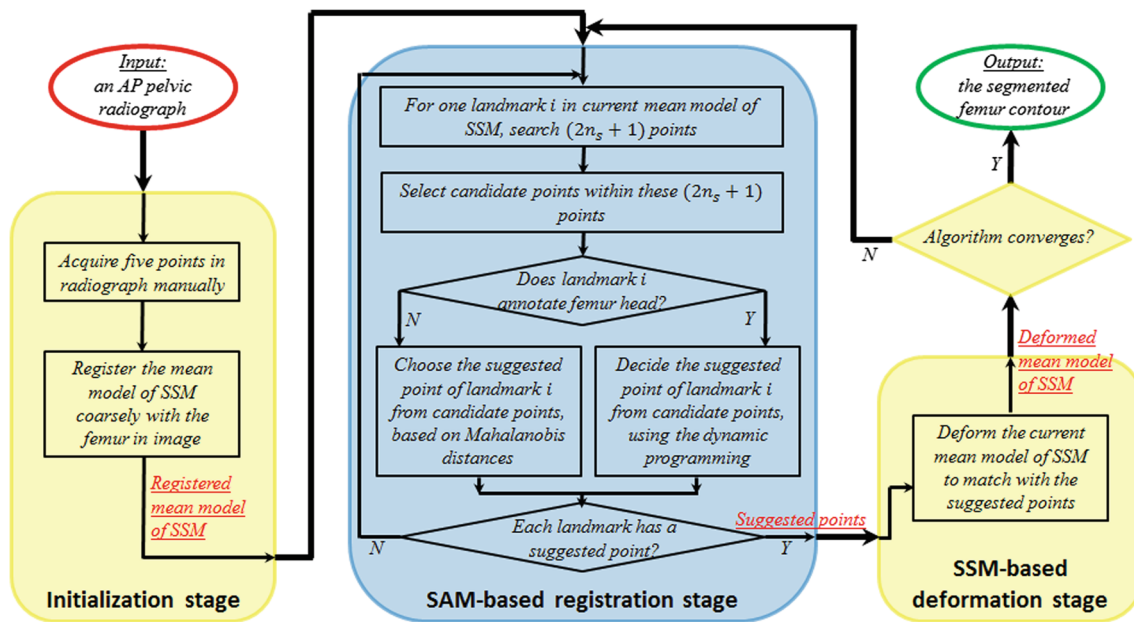


Fig. 5 Flowchart of the presented algorithm

programming technique [18] to determine its suggested point.

- Step 2-7: repeat steps 2-1 to 2-6, until every landmark on the mean model of the SSM is assigned with a suggested point.

3. SSM-based deformation stage

- Step 3-1: the suggested points selected/determined in the previous stage are considered as the next proposed positions of the corresponding landmarks.
- Step 3-2: apply a regularized non-rigid registration approach [1, 19, 20] to match the SSM with those proposed positions.

4. Repeat stages 2 and 3 until the algorithm converges.
5. Output the final SSM-registration result as the segmentation output.

Figure 5 shows the flowchart of the above algorithm. At the initialization stage, five points were manually picked in the given test image at positions of the greater trochanter, upper boundary contour point between femoral neck and head, fovea of femoral head, lower boundary point of neck and head and the lesser trochanter, respectively. These landmarks correspond to annotated landmarks of 21, 29, 37, 42 and 49, respectively. A scaled rigid registration algorithm [16, 17] was subsequently applied to match the mean shape of the SSM coarsely with the left proximal femur in the given test image. On-line supplemental Fig. IIa illustrates the five initially digitized points (red dots) and the original location of the mean model of the SSM (green stars and dots), while

on-line supplemental Fig. IIb shows the coarsely registered mean model (green stars and dots) after initialization.

Within the SAM-based registration stage, steps 2-1 to 2-5 were conducted for each landmark  $i$  on the mean shape of the SSM in order to search its suggested position that was the input to the SSM-based deformation stage. The Mahalanobis distance [14, 16] of a candidate point  $m$  was calculated as

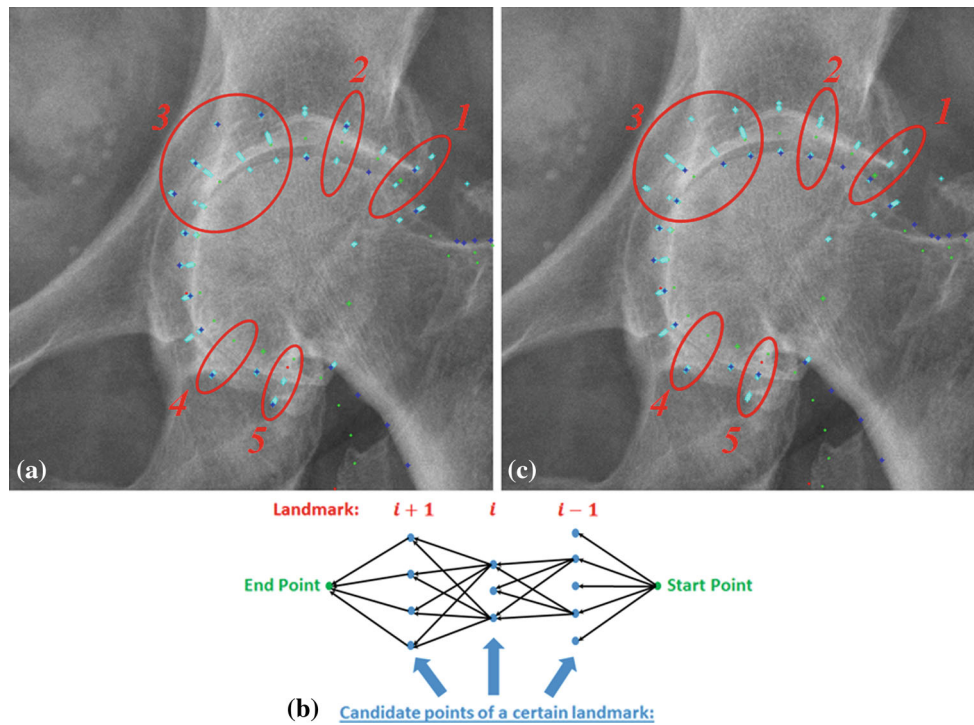
$$MD_m = \sum_{k=1}^K \frac{\alpha_k^2}{\lambda_k} \tag{3}$$

where  $\lambda_k$  is the  $k$ th largest eigenvalue of the SAM for the  $i$ th landmark associated with the candidate point  $m$ ,  $K$  is the number of largest eigenvalues of the SAM and

$$\alpha_k = (\mathbf{g}_m - \mathbf{g}_i) \cdot \mathbf{V}_k \tag{4}$$

where  $\mathbf{g}_m$  is the grey-level appearance model (i.e., the  $(2n_i + 1) \times 1$  derivative vector) of the candidate point  $m$ ,  $\mathbf{g}_i$  is the mean appearance model of the  $i$ th landmark and  $\mathbf{V}_k$  is the  $k$ th eigenvector corresponding to  $\lambda_k$ .

In the region around the femoral head, using only the minimal Mahalanobis distance to find the suggested points cannot achieve satisfactory results. Due to the complicated anatomical structures around this region, there are many candidate points having similar Mahalanobis distances. Therefore, the suggested point is hard to be properly chosen from those candidates. Red ellipses in Fig. 6a indicate such cases, where the cyan dots are the candidates and the blue stars are the suggested points improperly selected from these candidates for the associated landmarks.



**Fig. 6** **a** Examples show that the minimal Mahalanobis distance may not be a sufficient criterion for choosing the suggested points for landmarks in the femoral head region (enclosed by red ellipses). **b** Application of the dynamic programming technique to examples in (a). **c** Improved results after applying the dynamic programming technique

**Table 1** Statistics of segmentation errors in one hundred test radiographs (all in mm)

Results	Min	Max	Mean	Median	Std	25%-Error	75%-Error
Landmark segmentation error	0.00	16.46	0.95	0.40	1.57	0.20	1.00
Mean error	0.45	2.03	0.96	0.89	0.35	0.68	1.09

In order to address this problem, a dynamic programming-based technique [18] was used. For landmarks on the mean shape of the SSM that have been annotated as in the femoral head region, their suggested points were determined by minimizing the cost of paths passing through the candidate point of each landmark consecutively. The cost of a complete path was the sum of the Mahalanobis distances of all travelled candidates and the Euclidean distances between two consecutively passed-through candidates that were associated with two consecutive landmarks. Figure 6b illustrates the principle of the dynamic programming technique used in such a scenario. Figure 6c shows the improved results of the cases shown in Fig. 6a with the aid of the dynamic programming technique.

Results from the first iteration of SAM-based registration, namely the suggested points for all landmarks on the mean shape of the SSM are presented in on-line supplemental Fig. IIc.

In the SSM-based deformation stage, the suggested point for each landmark was input as the next proposed position of

the landmark on the mean shape of the SSM, a regularized non-rigid registration approach [1, 19, 20] was then applied to match the SSM to those proposed positions. This yielded the segmentation result (yellow dots and stars) as shown by the on-line supplemental Fig. II d.

Repeating the SAM- and SSM-based registration stages until the algorithm converges (empirically we found that a few iterations were needed), the segmentation result was obtained (see the on-line supplemental Fig. IIe for an example). B-spline interpolation was then used to get a continuous contour of the proximal femur (see the on-line supplemental Fig. II f for an example where the manually segmented femoral contour (green) serves as the ground truth for comparison).

**Results**

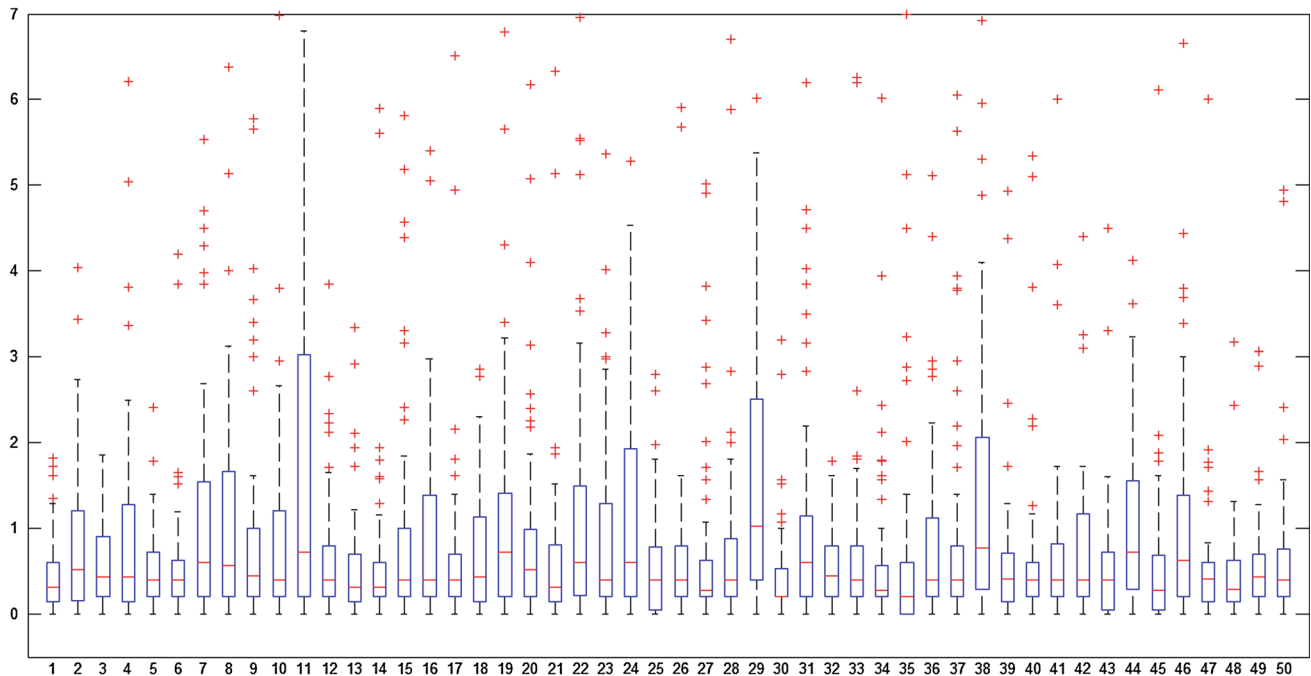
The proposed method was validated on one hundred clinical AP pelvic radiographs with a wide range of image qualities.

Please note that none of them were part of the training images.

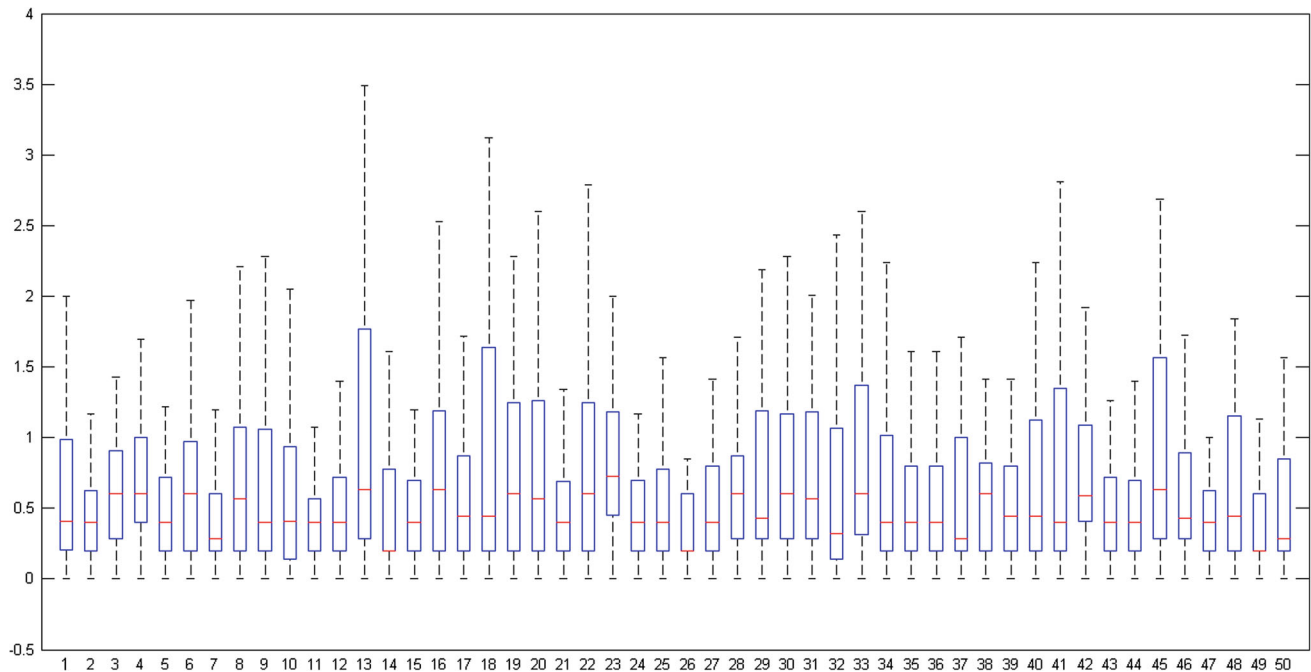
For each test radiograph, a manually segmented contour was used as the ground truth. The segmentation error was calculated as the point-to-curve distance, which is defined as the distance between each landmark obtained by the presented algorithm and the manually segmented contour. The

mean error of one test image was defined as the average segmentation error of all landmarks. Table 1 shows the errors when the results obtained with the presented algorithm were compared with the ground truth.

The boxplots of the segmentation errors of the first fifty test radiographs and the second fifty test radiographs are shown in Figs. 7 and 8, respectively. These boxplots were cre-

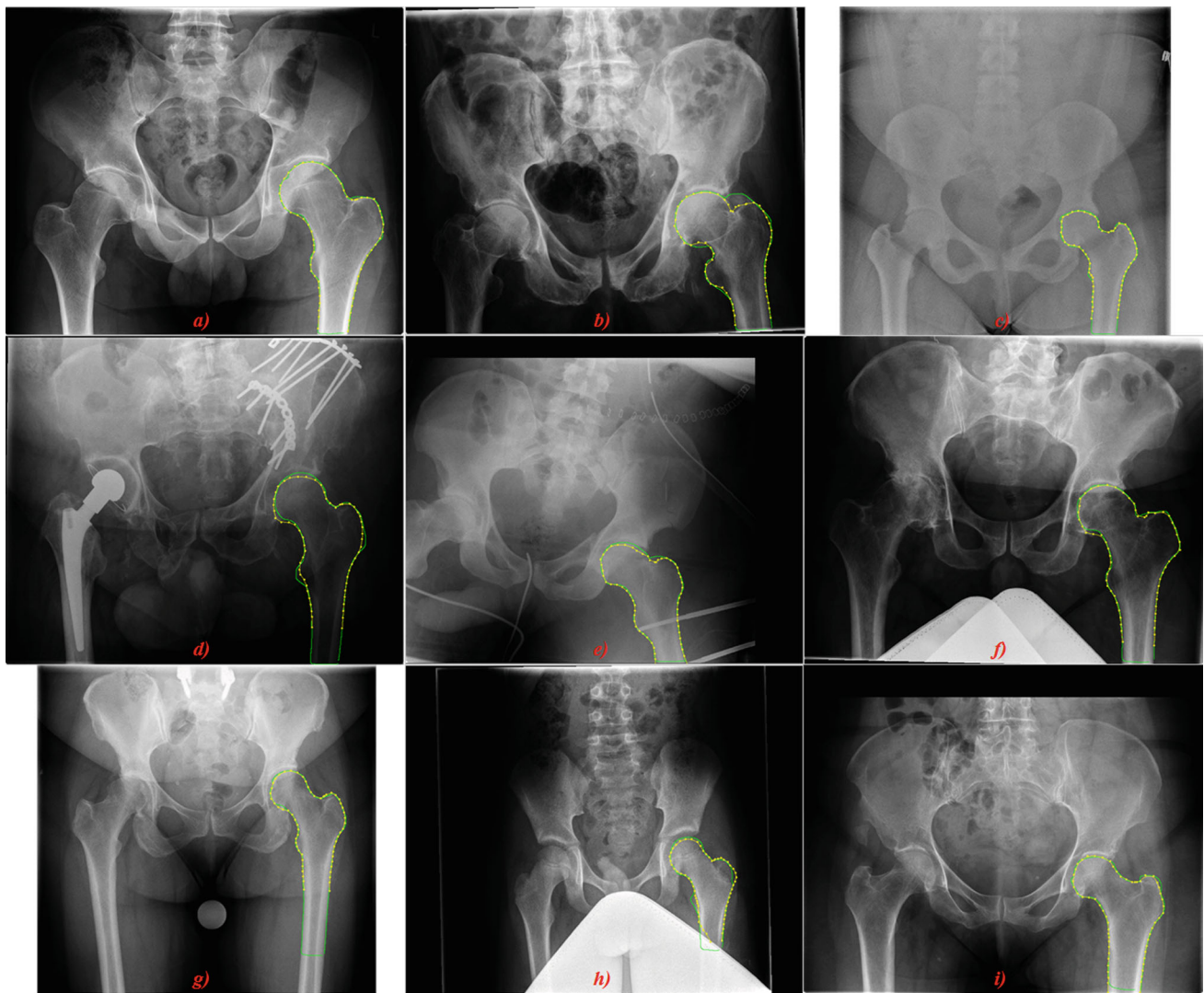


**Fig. 7** Boxplot of the segmentation errors of the first fifty test radiographs



**Fig. 8** Boxplot of the segmentation errors of the second fifty test radiographs





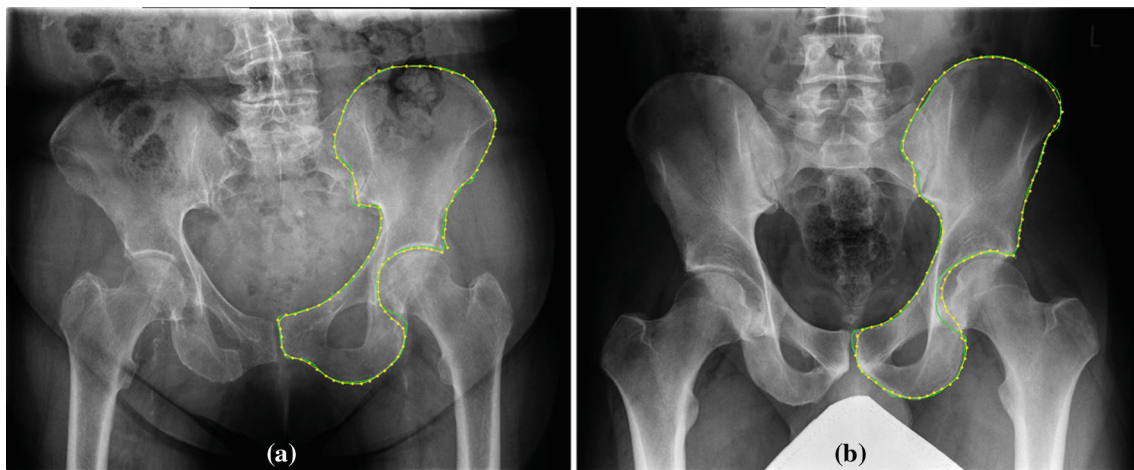
**Fig. 9** Examples of the interesting segmentation results (ground truth: green curve; result: yellow dots and contour) including: **a** the best segmentation result (mean error of 0.45 mm, with pathology of hip joint osteoarthritis); **b** the worst segmentation result (mean error of 2.03 mm, with pathology of hip joint osteoarthritis); **c** the segmentation result on an image with low quality due to much noise and poor contrast (mean error of 0.58 mm); **d** the segmentation result on an image with low quality due to weak contrast in the region of interest (mean error of 1.19 mm);

**e** the segmentation result on an image acquired in an abnormal pose (mean error of 0.91 mm, with pathology of hip joint osteoarthritis); **f** the segmentation result on an image of a hip joint with osteoarthritis (mean error of 0.64 mm); **g** the segmentation result on an image of a hip joint with dysplasia (mean error of 0.56 mm); **h** the segmentation result on an image of a child (mean error of 1.24 mm); **i** the segmentation result on an image with acetabulum fracture (mean error of 0.68 mm)

ated with Matlab<sup>®</sup> R2010b (The MathWorks, Inc., Massachusetts, USA). The horizontal axis denotes the number of the test radiograph, and the vertical axis is the registration error (in mm). On each box of both figures, the central magenta mark is the median, the upper and lower boundaries are the 25th and 75th percentiles, respectively, the whiskers extend to the maximum and the minimum, and the outliers are plotted with red symbols of “+” (Fig. 7).

Examples of interesting segmentation results are presented in Fig. 9, including the best segmentation result (a), the worst segmentation result (b), the segmentation result on an

image with low quality due to much noise and poor contrast (c), the segmentation result on an image with low quality due to weak contrast in the region of interest (d), the segmentation result on an image acquired in an abnormal pose (e), the segmentation result on an image with hip joint osteoarthritis (f), the segmentation result on an image with hip dysplasia (g), the segmentation result on an X-ray image of a child (h) and the segmentation result on an image with acetabulum fracture (i). These examples demonstrate the clinical applicability of the proposed approach. Implemented in an un-optimized Matlab environment, it requires approximately



**Fig. 10** Preliminary results of segmenting the left hemi-pelvis with the presented algorithm (ground truth: *green curve*; result: *yellow dots and contour*): **a** mean error 0.58 mm; **b** mean error 1.25 mm

**Table 2** Statistics of mean segmentation errors of twenty radiographs evaluated by invited clinicians (all in mm)

	Min	Max	Mean	Median	Std	25 %-Error	75 %-Error
Senior surgeon #1	0.47	1.75	1.04	0.92	0.46	0.63	1.49
Senior surgeon #2	0.54	2.48	1.19	0.93	0.54	0.78	1.52
Senior surgeon #3	0.52	2.23	1.05	0.88	0.47	0.73	1.37
Junior surgeon #1	0.50	2.03	1.09	0.82	0.53	0.65	1.47
Junior surgeon #2	0.57	2.46	1.03	0.86	0.50	0.69	1.07
Senior radiologist	0.47	2.02	1.09	1.00	0.45	0.68	1.31

five seconds per case for the presented algorithm to finish the segmentation.

In order to investigate the influence of the initialization on the segmentation results, six clinicians (three senior and two junior surgeons plus one senior radiologist) were invited to evaluate twenty out of the one hundred test radiographs used in our experiments. Without any training, they were requested to independently initialize the algorithm by manually digitizing five landmarks in each radiograph at positions of the greater trochanter, upper boundary contour point between femoral neck and head, fovea of femoral head, lower boundary point of neck and head, and the lesser trochanter, respectively. Based on their initializations, the algorithm automatically computed the segmentation results. Table 2 shows the mean errors (all in mm) of the segmentation results. With the significant level  $\alpha = 0.05$ ,  $t$ -test was used to determine whether the differences between segmentation results from the six clinicians and those from the one hundred test radiographs were statistically significant. All results were analysed with a MATLAB function developed by us (for calculation of  $t$ -values) and a free statistics calculator (for calculation of  $p$  values) (<http://www.danielsooper.com/statcalc3/calc.aspx?id=8>). The respective two-tailed  $p$  values of the segmentation results of each clinician were 0.46, 0.07, 0.42,

0.29, 0.55 and 0.22, when they are compared to those from one hundred test radiographs. Thus, it was found that the differences between the measurement results from the six clinicians and those from the one hundred test radiographs were not statistically significant. Therefore, a conclusion could be drawn that statistically, there is no significance difference in terms of segmentation accuracy when initializations from different end users are used.

## Discussion and conclusions

A new method for segmenting the proximal femur in digital AP pelvic X-ray radiographs was proposed, which combined the statistical shape and appearance models with the dynamic programming technique.

In the method of Chen et al. [10], unexpected errors were introduced because spurious edges detected in images might influence the initialization of the initial 2D model. Similarly, in that of Boukala et al. [11], spurious edges existing in images might lead to an incorrect counting of edge points in shape context descriptor, and thus, the inappropriate correspondences were assigned, which resulted in the failed initialization of the initial mean shape of SSM. Unlike methods

in Chen et al. [10] and Boukala et al. [11], the proposed approach initialized the mean model of SSM by manually acquiring five points in radiograph, rather than by searching parallel lines and circles in femur region or by explicitly finding correspondence with shape context descriptor. Therefore, it was not affected by spurious edges detected in images, whereas above both methods did.

The method presented by Ding et al. [13] required an atlas describing not only femurs but also the pelvis in order to make the global alignment robust against spurious edges. Thus, it may not work properly on those standard AP radiographs for THA where only part of the pelvis is imaged to be visible (see the on-line supplemental Fig. 1d for an example). Different from the method introduced by Ding et al. [13], the proposed approach did not require an atlas of the pelvis to achieve a robust initialization against spurious edges, and thus, it is applicable to those standard AP radiographs for THA where only part of the pelvis is visible.

Experiments carried out on one hundred clinical AP pelvic radiographs (in a wide range of image qualities) showed an average mean error of 0.96 mm and a standard deviation of 0.35 mm. It was more accurate than those in Behiels et al. [7] and Pilgram et al. [8], where both works did not reach very high accuracy (with mean segmentation error of approximately 1.9 mm). Due to the combination of both SSM and SAM with the dynamic programming technique, the proposed approach achieved a high accuracy that is comparable to that reported by Lindner et al. [9]. The method in Lindner et al. [9] did not take into account segmentation of the lesser trochanter, whereas the proposed approach did. In certain clinical applications, for instance, transtrochanteric valgus osteotomy (TVO) [21] and risk prediction of hip fracture [22], a 3D patient-specific femur model will be a helpful tool to assist surgical planning and validate effectiveness of different diagnosis criteria. However, the reconstruction [1, 2] of such a model requires the extraction of the complete femoral contour, including the lesser trochanter.

Compared with the method described by Gamage et al. [12], the proposed method retained a higher performance, even in case when radiographs were acquired in an abnormal pose. In addition, the segmentation results reported in Gamage et al. [12] were less convincing since their method was validated with a small number (fifteen) of test cases.

In conclusion, we developed and tested a robust, accurate and efficient algorithm for segmenting the proximal femur in AP pelvic X-ray radiographs. Our method is general enough to be extended to segmenting other bony structures, such as the pelvis (see Fig. 10 for an example) in an AP pelvic X-ray radiograph.

**Acknowledgments** This work was partially supported by the Swiss National Science foundation via project No. 205321\_138009.

**Conflict of interest** Weiguo Xie, Jochen Franke, Cheng Chen, Paul A. Grützner, Steffen Schumann, Lutz-P. Nolte, and Guoyan Zheng declare that they have no conflict of interest.

## References

- Zheng G (2010) Statistical shape model-based reconstruction of a scaled, patient specific surface model of the pelvis from a single standard AP X-ray radiograph. *Med Phys* 37(4):1424–1439
- Zheng G (2010) Statistically deformable 2D/3D registration for estimating post-operative cup orientation from a single standard AP X-ray radiograph. *Ann Biomed Eng* 38(9):2910–2927
- Wright D, Whyne C, Hardisty M, Kreder HJ, Lubovsky O (2011) Functional and anatomic orientation of the femoral head. *Clin Orthop Relat Res* 469:2583–2589. doi:10.1007/s11999-010-1754-1
- Kellgren JH, Lawrence JS (1957) Radiological assessment of osteoarthritis. *Ann Rheum Dis* 16:494–501
- Turmezei TD, Poole KES (2011) Computed tomography of subchondral bone and osteophytes in hip osteoarthritis: the shape of things to come? *Front Endocrinol* 2:97. doi:10.3389/fendo.2011.00097
- Charbonnier C, Magnenat-Thalmann N, Becker CD, Hoffmeyer P, Menetrey J (2010) An integrated platform for hip joint osteoarthritis analysis: design, implementation and results. *Int J Comput Assist Radiol Surg* 5(4):351–358
- Behiels G, Vandermeulen D, Maes F, Suetens P, Dewaele P (1999) Active shape model-based segmentation of digital X-ray images. In: Taylor C et al (eds) MICCAI 1999, LNCS 1679. Springer, Heidelberg, pp 128–137
- Pilgram R, Walch C, Kuhn V, Schubert R, Staudinger R (2008) Proximal femur segmentation in conventional pelvic X ray. *Med Phys* 35(6):2463–2472
- Lindner C, Thiagarajah S, Wilkinson JM et al (2012) Accurate fully automatic femur segmentation in pelvic radiographs using regression voting. In: Ayache N et al (eds) MICCAI 2012, Part III, LNCS 7512. Springer, Heidelberg, pp 353–360
- Chen Y, Ee X, Leow WK, Howe TS (2005) Automatic extraction of femur contours from hip X-ray images. In: Liu Y et al (eds) ICCV Workshop on CVBIA 2005, LNCS 3765. Springer, Heidelberg, pp 200–209
- Boukala N, Favier E, Laget B, Radeva P (2004) Active shape model based segmentation of bone structures in hip radiographs. In: IEEE international conference on industrial technology (ICIT'04), pp 1682–1687. doi:10.1109/ICIT.2004.1490821
- Gamage P, Xie SQ, Delmas P, Xu WL (2010) Segmentation of radiographic images under topological constraints: application to the femur. *Int J Comput Assist Radiol Surg* 5:425–435. doi:10.1007/s11548-009-0399-6
- Ding F, Leow WK, Howe TS (2007) Automatic segmentation of femur bones in anterior-posterior pelvis X-Ray images. In: Kropatsch WG et al (eds) CAIP 2007, LNCS 4673. Springer, Heidelberg, pp 205–212
- Cootes TF, Hill A, Taylor CJ, Haslam J (1994) The use of active shape models for locating structures in medical images. *Image Vis Comput* 12(6):355–366
- Cootes TF (2000) An introduction to active shape models. In: Baldock R, Graham J (eds) Image processing and analysis. Oxford University Press, Oxford, pp 223–248
- Xie W, Schumann S, Franke J, Grützner PA, Nolte L-P, Zheng G (2012) Finding deformable shapes by correspondence-free instantiation and registration of statistical shape models. In: Wang F et al

- (eds) MICCAI Workshop on MLMI 2012, LNCS 7588. Springer, Heidelberg, pp 258–265
17. Xie W, Nolte LP, Zheng G (2011) ECM versus ICP for point registration. In: 33rd Annual international conference of the IEEE engineering in medicine and biology society (EMBC'11), pp 2131–2135. doi:[10.1109/IEMBS.2011.6090398](https://doi.org/10.1109/IEMBS.2011.6090398)
  18. Bellman R (1957) Dynamic programming. Princeton University Press, Dover paperback edition (2003), ISBN 0-486-42809-5
  19. Zheng G, Gollmer S, Schumann S, Dong X, Feilkas T, Ballester MAG (2009) A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. *Med Image Anal* 13(6):883–899
  20. Zheng G, von Recum J, Nolte L-P, Gruetzner PA, Steppacher SD, Franke J (2012) Validation of a statistical shape model-based 2D/3D reconstruction method for determination of cup orientation after THA. *Int J Comput Assist Radiol Surg* 7(2):225–231
  21. Jingushi S, Sugioka Y, Noguchi Y, Miura H, Iwamoto Y (2002) Transtrochanteric valgus osteotomy for the treatment of osteoarthritis of the hip secondary to acetabular dysplasia. *J Bone Joint Surg Br* 84-B(4):535–539
  22. Li W, Kornak J, Harris T et al (2009) Identify fracture-critical regions inside the proximal femur using statistical parametric mapping. *Bone* 44(4):596–602. doi:[10.1016/j.bone.2008.12.008](https://doi.org/10.1016/j.bone.2008.12.008)