Cascaded statistical shape model based segmentation of the full lower limb in CT

Emmanuel A. Audenaert, Jan Van Houcke, Diogo F. Almeida, Lena Paelinck, M. Peiffer, Gunther Steenackers & Dirk Vandermeulen

Published online: 01 Mar 2019.

Submit your article to this journal

Article views: 978

View related articles

View Crossmark data

Citing articles: 34 View citing articles

Taylor & Francis
Taylor & Francis Group

Check for updates

Cascaded statistical shape model based segmentation of the full lower limb in CT

Emmanuel A. Audenaert[a,b,c], Jan Van Houcke[a] (iD), Diogo F. Almeida[a], Lena Paelinck[a], M. Peiffer[a], Gunther Steenackers[c] and Dirk Vandermeulen[d]

[a]Department of Orthopedic Surgery and Traumatology, Ghent University Hospital, Ghent, Belgium; [b]Department of Trauma and Orthopedics, Addenbrooke's Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK; [c]Department of Electromechanics, Op3Mech research group, University of Antwerp, Antwerp, Belgium; [d]Department of Electrical Engineering, ESAT/PSI, KU Leuven, Leuven, Belgium

ABSTRACT

Image segmentation has become an important tool in orthopedic and biomechanical research. However, it greatly remains a time-consuming and laborious task. In this manuscript, we propose a fully automatic model-based segmentation pipeline for the full lower limb in computed tomography (CT) images. The method relies on prior shape model fitting, followed by a gradient-defined free from deformation. The technique allows for the generation of anatomically corresponding surface meshes, which can subsequently be applied in anatomical and mechanical simulation studies. Starting from an initial, small ($n \leq 10$) sample of manual segmentations, the model is continuously updated and refined with newly segmented training samples. Validation of the segmentation pipeline was performed by comparing the automatic segmentations against corresponding manual segmentations. Convergence of the segmentation pipeline was obtained in 250 cases and failed in three samples. The average distance error ranged from 0.53 to 0.76 mm and maximal error ranged from 2.0 to 7.8 mm for the 7 different osteological structures that were investigated. The accuracy of the shape model-based segmentation gradually increased as the number of training shapes in the updated population also increased. When optimized with the free form deformation, however, average segmentation accuracy rapidly plateaued from already as little as 20 training samples on. The maximum segmentation error plateaued from 100 training samples on.

Abbreviations

SSM	Statistical Shape Model;
CT	Computed Tomography;
G-RBF	Gaussian Radial Basis Function;
ICP	Iterative Closest Point;
NN	Nearest Neighbor;
SVD	Singular Value Decomposition;
GPA	General Procrustes Alignment;
PCA	Principal Component Analysis;
ROI	Region of Interest;
RMSE	Root Mean Squared Error;
ME	Maximum Error

1. Introduction

Statistical shape and intensity models (SSIM) have been established as a robust tool for the segmentation of medical images and analysis of variation in anatomy (Cootes et al. 1994, 1995; Almeida et al. 2016). Due to their ability to construct detailed personalized models from sparse data sets, they have the potential to efficiently generate patient-specific models. This permits the generation of large, simulated populations for the use in injury biomechanics, musculoskeletal disease models or implant design optimization (Baldwin et al. 2010; Henak et al. 2013).

While SSIM have become popular for the description of isolated anatomical structures, they have rarely been applied in more complex integrated anatomies such as an entire limb or even the entire human skeleton. Training such models would require a large number of samples to be available and be pre-processed (Henak et al. 2013). In order to build a SSIM, the first step consists in segmenting the volumes of interest in a reasonably representative amount of training data. Manually assisted segmentation of a full articulated limb, however, is a time consuming and arduous task, which has been estimated to range over 100 man-hours per case (Henak et al. 2013). Data

samples including over 2000 slices (each consisting of 512*512 pixels) per subject are not uncommon and the massive size of this data type involves an additional computational bottleneck for several (semi)-automatic segmentation routines.

In the present paper we present a generic, cascaded approach for bone segmentation, that can initialize from minimal training data. The method relies on articulated. multi-object, SSM based fitting, followed by single-object SSM fitting and finally, gradient-defined free from deformation in combination with digital subtraction of gradient peaks of pre-segmented neighboring structures to avoid overlap around joints. Segmentation accuracy evolution was evaluated with increasing training samples being added to the SSM. Understanding the impact of a growing amount of SSM training samples can guide future developers by assuring sufficient training samples for either segmentation purposes either population wide applications of SSMs.

Related work

1.1. Articulated shape models

Describing population variance in terms of shape of isolated objects is a challenging job requiring substantial training samples to achieve acceptable generalization properties of the models. While information on required number of training samples for covering the variance of skeletal structures in a given population is scarce in the literature, it has been documented in the area of face recognition, where studies have shown that in order to representatively achieve a population covering study, at least 250 samples are required (Claes 2007). Taking in account that for musculoskeletal applications we need to describe the shape of connected structures presenting as well as their relative positioning anatomical alignment (e.g. knee varus/valgus alignment), the dimension of our shape space is expected to be even higher. Different poses among scanned patients at the time of the image acquisition need to be taken into account which contributes to the complexity of an already challenging problem. Previous authors have attempted to reduce the computational impact of alignment and positional variance for segmentation applications, either by defining statistical transformation models to capture differences in pose (Boisvert et al. 2008) or by explicitly modeling joint motion trough idealized joint models (e.g. spherical in case of the hip or hinged for the knee joint) (Kainmueller et al. 2008, 2009).

1.2. Segmentation of bone and joints

The main challenge in joint segmentation lies in the close spatial relationship of the objects to be segmented, potentially causing penetration of surfaces and overlap. This is particularly challenging in osteo-arthritic cases, where the wearing down of cartilage causes bone-on-bone contact, that makes articulating bones almost impossible to distinguish. In addition, osteoporotic bones translate as weak boundaries in CT-Scans which makes thresholding and edge-based techniques fail at accurately segmenting the bone geometries (Bourgeat et al. 2007). Graph-based algorithms have also been extensively explored for the purpose of bone segmentation, but the accuracy of these algorithms often relies on seed points manually provided (Boykov and Funka-Lea 2006; Shim et al. 2009). Furthermore, the bones are usually segmented individually instead of jointly, which can lead to sub-optimal segmentation results in terms of segmentation overlap in joint areas. To handle the bone overlap, Li et al. proposed a novel column graph-based algorithm to solve coupled surface segmentation problems, which was later used for simultaneous bone and cartilage segmentation in the knee (Li et al. 2006; Yin et al. 2010).

Atlas-based and statistical shape model based methods are an alternative approach commonly used to rely on prior knowledge to guide and improve the segmentation (Chu et al. 2015). As segmentation of single anatomical structures doesn't consider the relationship with neighboring structures, Kainmueller et al. introduced coupled or articulated deformable models for multiple-object segmentation, which do not completely prohibit but discourage the overlap (Kainmueller et al. 2009; Bindernagel 2013). To further avoid such overlap Yokota et al. suggested a cascaded approach by excluding edges of pre-segmented regions in the SSM based fitting procedure (Yokota et al. 2009, 2013).

In recent years, significant developments have been made for image processing and analysis as a result of the revival of deep convolutional neural networks (CNNs), both in terms of accuracy and processing speeds once trained (Gassman et al. 2008; Liu et al. 2018). To date and to the authors knowledge, the main issue is the data size of the input layers and the availability of large sets of labeled training data. Several data augmentation techniques have been described to partially deal with these issues and established pre-trained networks, such as the U-net, have been made public (Ronneberger et al. 2015). CNNs are versatile in their application domain, going way

beyond image segmentation, as they can be trained for 2D to 3D registration purposes, disease recognition and quantification (Antony et al. 2016). While such techniques are currently still being explored and developed, approaches such as SSIM segmentation techniques stand as valid alternatives as the computational cost for training is significantly lower and they are able to provide the necessary amount of labeled training data for advanced neural network applications (Bhalodia et al. 2018).

2. Methods

A statistical shape model-based approach that is initiated from a minimal sample of manual segmentations ($n \leq 10$) was implemented and the shape model was gradually updated and refined as more segmentations became available. The following osteological structures were segmented: D12-L5 vertebrae, sacrum, pelvis, femur, patella, tibia, fibula, talus, calcaneum, navicular, cuboid and cuneiform bones.

In Section 2.1.1, the surface registration procedure is described. It is used to define point correspondences as required for the construction of the different statistical shape models (Section 2.1.2). The automatic segmentation pipeline, fitting the previously built SSM to new image data, is described in Section 2.2. The validation procedure for the complete segmentation pipeline is presented in Section 2.3.

2.1. Construction of the initial deformable SSM models

2.1.1. Dense landmarking of manual segmentations
An iterative closest point (ICP) based non-rigid registration algorithm was used to register template meshes of all osteological segments, further referred to as source meshes, to a small set ($n \leq 10$) of manually segmented cases, further referred to as target meshes. The source mesh was created by isotropic remeshing a randomly selected manually segmented case. First, target and source meshes, respectively $\{v_i^t, \ i \in 1, ..., N^t\}$ and $\{v_i^s, \ i \in 1, ..., N^s\}$, are aligned according to their principal axes of inertia (Esat and Bahai 2000). Then, the rigid alignment is optimized with an adapted version of the rigid ICP algorithm, using bidirectional (target to source and back) vertex correspondences and excluding distance-based outliers (Besl and McKay 1992).

The forward vertex correspondence $v_{\rightarrow}^t(i)$ is defined as the nearest neighbor (NN) of the source vertex v_i^s on the target surface:

$$v_{\rightarrow}^t(i) = \ \text{argmin}_{v_l^t}\left(\left\|v_l^t - v_i^s\right\| \ \mid \ l \in 1, ..., N^t\right) \quad (1)$$

The backward vertex correspondences $v_{\leftarrow}^t(i)$ are defined as the vertices v_l^t on the target surface for which v_i^s is their NN.

$$v_{\leftarrow}^t(i)$$

$$= \ \left\{v_l^t \mid v_i^s = \text{argmin}_{v_j^s}\left(\left\|v_j^s - v_l^t\right\| \ j \in 1, ..., N^s, l \in 1, ..., N^t\right)\right\} \quad (2)$$

For every source vertex v_i^s, there is a set of displacements $\delta(v_i^s)$, one for the forward $\delta_{\rightarrow}(.)$ and one or more for the backward $\delta_{\leftarrow}(.)$ correspondences:

$$\delta_{\rightarrow}(v_i^s) = v_{\rightarrow}^t(i) - v_i^s \quad (3)$$

and

$$\delta_{\leftarrow}(v_i^s) = v_l^t - v_i^s, \ \forall \ v_l^t \in v_{\leftarrow}^t(i) \quad (4)$$

The additional backward displacements are introduced to warp the source surface into strong indentations of the target surface.

Given these vector displacements, first, the optimal rotation ($\check{R}$), translation ($\check{t}$) and scaling ($\check{s}$) are calculated iteratively as solutions to the following weighted least squares minimization problem:

$$T\{\check{R}, \check{t}, \check{s}\} = \ \text{argmin}_{R,t,s}\left(\sum_{i=1}^{N^s} w_i \left\|T(v_i^s) - (v_i^s + \delta_{\rightarrow}(v_i^s))\right\|^2\right.$$

$$\left. + \sum_{l=1}^{N^t} w_l \left\|T(v_l^s) - (v_l^s + \delta_{\leftarrow}(v_i^s))\right\|^2\right) \quad (5)$$

with $T(v) = sRv + t$ the similarity transformation and w_i a binary indicator variable coding for inliers ($w_i = 1$) and outliers ($w_i = 0$). The inner summation is shorthand notation for summing over all elements of $\delta(v_i^s)$. For every iteration, a normalized histogram of $\{\left\|\delta(v_i^s)\right\|, i = 1, ..., N^s\}$ is calculated, and vertices v_i^s are considered as outliers if $\left\|\delta(v_i^s)\right\|$ is in the top-α percent range (typically set at 5%) and consequently discarded. At each iteration, $\check{R}, \check{t}$ and $\check{s}$ are calculated as solutions of a Singular Value Decomposition (SVD) (Arun et al. 1987) and new displacements $\delta(v_i^s)$ are determined. The algorithm converges if the root mean squared error between the previous and current source vertices drops below a predefined threshold value ($< 10^{-5}$mm).

After rigidly aligning the resulting vector displacements and mapping source vertices to the target mesh, non-rigid registration of the source mesh onto the target mesh is applied. In a first step we approximate remaining vector displacements mapping source vertices to the target mesh using a non-rigid transformation, modelled as a sum of N_c Gaussian Radial Basis Functions (G-RBF) with variable γ defining the width of the Gaussian function (Carr et al. 2001; Chui and Rangarajan 2003; de Boer et al. 2007; Kim et al. 2016):

$$d(\cdot) = \sum_{k=1}^{N_c} w_k \varphi(\cdot - \mathbf{c}_k) \text{ with } \varphi = e^{-(\gamma(\|\cdot\|))^2} \qquad (6)$$

satisfying the following constraints:

$$\delta(v_i^s) = \sum_{k=1}^{N_c} w_k \varphi(v_i^s - \mathbf{c}_k), \ \ i = 1, ..., N^s \qquad (7)$$

which is shorthand for satisfying the constraint for every element of the set $\delta(v_i^s)$. The G-RBF centers $\mathbf{c}_k$ are, ideally, located at the source vertices v_i^s. For computational efficiency, the number N_c of G-RBF centers was chosen to be a much smaller subset of the source vertices. As a result, the deformation defined at the G-RBF centers is interpolated at all other source vertices. This turns the set of $3M$ equations (M counting the sum of the cardinalities of $\delta(v_i^s)$. In the N_c 3-D vector unknowns w_k into an overdetermined problem, which was solved in a least squares sense by minimizing:

$$\sum_i^{N^s} \left\| d(v_i^s) - \delta(v_i^s) \right\|^2 + \lambda \|\mathbf{c}\|^2 \qquad (8)$$

with an additional Tikhonov L_2-regularization term $\lambda \|\mathbf{c}\|^2$ added, preventing instability/ill-conditioning of the system with the internal summation summing over all elements of $\delta(v_i^s)$. (Tikhonov and Arsenin 1977).

This ultimately results in the following set of (regularized) normal equations for the x-, y- and z-components of the deformation coefficients $w_\cdot$, given the displacements $\delta(v_\cdot)$:

$$w_\cdot = (\Phi^T \Phi + \lambda I)^{-1} \Phi^T \ \delta(v_\cdot) \text{ with } \phi_{i,k} = \ \varphi(v_i^s - \mathbf{c}_k) \qquad (9)$$

Note that, according to the definition of the displacements (Eqs. (3) and (4)), we have one or several equations per source vertex. This framework for determining the vertex correspondences and the approximating G-RBF transformation is solved iteratively using a deterministic annealing scheme. Decreasing deformation stiffness is hereby obtained by

Table 1. Overview of the mesh properties of the different segments investigated: number of vertices, number of faces and average length of the triangle edges.

	Vertices	Faces	Edge length (mm)
Pelvis	103872	207736	1.57 ± 0.27
Femur	42194	84376	1.77 ± 0.31
Tibia and Fibula	78394	156772	1.24 ± 0.32
Patella	7650	15292	1.05 ± 0.21
Calcaneum and Talus	19644	39272	1.40 ± 0.27
Lumbar 5	12564	25128	1.55 ± 0.32
Sacrum	32610	65348	1.70 ± 0.50
Midfoot bones	20662	41276	1.38 ± 0.34
Lumbar Spine (D12-L5)	75384	150768	1.48 ± 0.48
Full lower limb	58113	116258	4.75 ± 2.92

progressively increasing the number N_c of G-RBF centres while decreasing the G-RBF width defined by γ.

While this approach generates a smooth warp from source to target mesh, it remains constrained, in terms of number of degrees of freedom, to the number of G-RBF centers defined and, as a result, can provide locally suboptimal results. This could be resolved by increasing the number of G-RBF centres to the number of source vertices. However, the computational load for typical mesh sizes (see Table 1) is far too high, still. Instead, the previous warp was refined by a local elastic deformation, defined as a weighted locally rigid transformation aligning source and target mesh (Li et al. 2008). For each source vertex v_i^s, a local neighbourhood $\aleph_i$ of vertices on the source mesh is defined, based on spatial (Euclidean distance) and orientation (normal direction) similarity. The local rigid deformation (rotation matrix R_i and translation vector t_i) mapping these source points to their nearest neighbours on the target mesh is then solved in a least squares sense (similar to Eq. (5)) by SVD.

Next, a weighted rigid transformation is applied to the whole mesh (Eqs. (10) and (11)).

$$l(v_i^s) = \sum_{j \varepsilon \aleph_i} w_{ji} \left[R_j(v_i^s) + t_j(v_i^s) \right] \qquad (10)$$

with linear distance based weights

$$w_{ji} = \frac{1 - d_{ji}/\overline{d}}{(N-1)}, \quad \sum_j^N w_{ji} = 1, \ \ d_{ji} = \ \ \| v_j^s - v_i^s \|,$$

$$\overline{d} = \sum_{l \in \aleph_i} (\| v_l^s - v_i^s \|)/N \text{ and } N = (\aleph_i). \qquad (11)$$

2.1.2. PCA of aligned corresponding shapes

The procedure of template/source to mesh registration, outlined in the previous section, generated sets of meshes with homologous vertices based on the initial manual segmentations available.
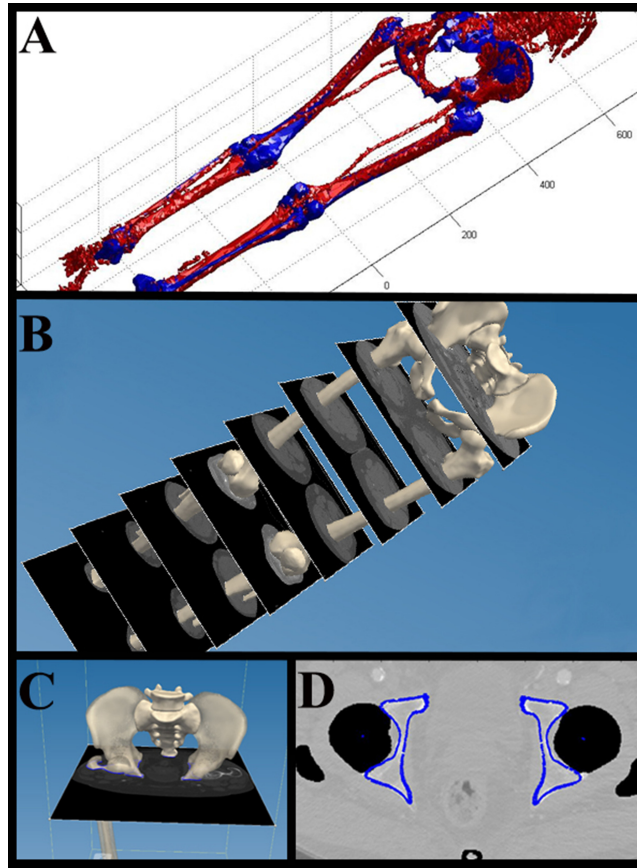
Figure 1. Overview of the segmentation pipeline: Initial positioning and fitting of the low-resolution shape model (blue) on the isosurface reconstruction of the CT scan (red) (A), fitting of the full lower limb shape model to the image (B), followed by in-image high resolution statistical shape fitting per segment (C) and free-form high-resolution registration in image per segment (D).

Generalized Procrustes Alignment (GPA) was applied to align the homologous sets, removing similarity (rotation, translation, scaling) transformations differences, prior to applying Principal Component Analysis (PCA) to determine the (co-)variance of morphological differences within the data set.

As a result, a statistical shape model (Cootes et al. 1995) was generated, described as

$$S = \overline{S} + Pb \qquad (12)$$

with S the shape vector represented as the ordered list of vertex coordinates (following GPA alignment). $\overline{S}$ defines the corresponding average shape, $P = (p_1, p_2, ...p_t)$ the matrix of eigenvectors of the covariance matrix $(S-\overline{S})^T (S - \overline{S})$, and $b = (b_1, b_2, ...b_t)^T$ a vector of weights. The eigenvectors are associated and ranked in decreasing order according to the t largest eigenvalues, each representing a mode of shape variation as observed over the training data set. The number t of eigenvectors retained is determined such that the cumulative variance represented is 98% of the total variance observed. This shape model is further updated as new, verified segmentations (section 2.2) become available.

Different segmental shape models were constructed, as well as an articulated full lower limb model of lower resolution (obtained by geometry preserving patch reduction). An overview of mesh properties of the different mean shape models is provided in Table 1.

2.2. Segmentation pipeline

With the focus on fully automatic segmentation, our method follows a global-to-local approach divided into four steps.

1. Initial fitting of the average shape of the full-lower limb model to the target mesh to estimate position, orientation and scaling of the full lower limb. The target mesh is obtained as the triangulation of a thresholded isosurface and represents a crude initial segmentation. Subsequently the
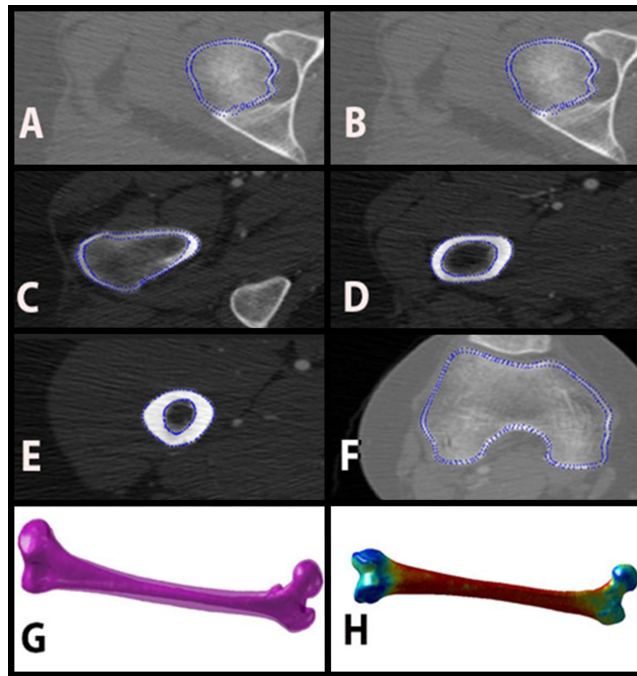
Figure 2. Detail of the femur in-image segmentation process at high resolution (A to F) with inner and outer cortical shape (G) and density (H) extraction.

statistical low-resolution shape model is fitted to the triangulated target mesh.

2. Statistical (in-image) shape fitting of the low-resolution articulated lower limb model to the target image.

3. Statistical (in-image) shape fitting of each single-object high-resolution SSMs to the target image.

4. Free-form (in-image) segmentation of each separate segment. This final step is required to reduce the restrictive character of SSM and to extend the shape model to include verified morphology not yet captured by the SSM.

The fully automatic 3D segmentation pipeline is summarized in Figure 1.

2.2.1. Initialization

The first step starts with a fourfold down-sampling of the CT-volume in order to reduce the calculation time for the initialization. The image is segmented using automatic Otsu-thresholding with two thresholds (Otsu 1979), roughly separating the volume of interest from the background, respectively bone and soft tissue. Secondly, the binary dataset corresponding to the cortical bone is corrected using a series of 3D binary morphological operations. The specific order of these operations was: bridging unconnected pixels, closing open regions and subsequently filling of holes. Unconnected parts smaller than a predefined 100 pixels and evaluated at a connectivity of 26 were than

considered noise and removed. Following, the outer perimeter pixels were calculated. The isosurface generated from this volume data provides a first estimate of the geometry of the cortical surfaces in the target image (Figure 1A – red surface).

As a first estimate of shape and position of the lower limb bones, the average shape of the low-resolution shape model of the lower limb is rigidly fitted onto the cortical isosurface using a iterative closest point (ICP) approach. Similar to the registration process described in Section 2.1.1, the fitting procedure iteratively defines bidirectional correspondences and excludes outlier candidate points, which typically occur at interruption sites of the meshes or are due to local noise. Iterative solutions of the rotation, translation and scaling parameters are obtained by SVD. The algorithm converges if the root mean squared error between the previous and current shape model vertices drops below a predefined threshold value ($<10^{-5}$ mm) (Figure 1A – blue surface, fitted low-resolution lower-limb shape model).

Secondly, the lower limb statistical shape model, constructed in Section 2.1.2, is fitted with increasing degrees of freedom (gradually increasing the modelling flexibility by increasing the number t of principal components retained in Eq. (13)) onto the cortical isosurface by iteratively solving an overdetermined ($N^s \gg t$) set of linear regression equations
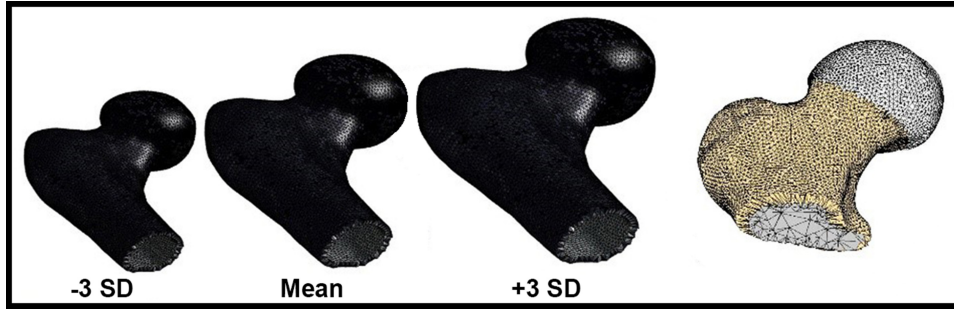
Figure 3. Generated solid cortical, cartilage and marrow volumes based on the correspondence feature of the inner and outer cortical mesh layer. Mean proximal femur and +/-3 standard deviations of the first principal component describing size are presented (left). Solid cartilage layers were added (right) by applying a distance offset along the surface normal of joint.

in a least squares sense, while adapting for pose (rotation, translation, scale) changes $T(.)$ using SVD.

$$\tau + \delta\tau \overset{\text{lsq}}{\cong} T\left(\overline{S} + \sum_{i=1}^{t} p_i b_i\right) \tag{13}$$

$$b = P^{\mathrm{T}}\left(T^{-1}(\tau + \delta\tau) - \overline{S})\right) \tag{14}$$

[Note: since the vectors p_i are orthonormal, $(P^{\mathrm{T}}P)^{-1} = I$ and the equation $b = (P^{\mathrm{T}}P)^{-1}P^{\mathrm{T}}\ \delta(v)$ can be simplified.]

with τ the current shape model vertex positions and $\delta\tau$ the current displacement vector to the targeted solution (Eq. (14)). The goodness-of-fit criterion for convergence is defined by the sum of squared errors ($<10^{-5}$).

2.2.2. In-image fitting of the low resolution articulated lower limb SSM

In the second step, the fitting result of the low-resolution lower limb SSM to the initial isosurface is further improved by fitting the articulated low-resolution SSM directly to the image data (in-image fitting, Figure 1B). Again, an ICP fitting procedure is used, this time maximizing the total sum of target image gradients (taking gradient direction versus surface normal into account) evaluated at the template mesh vertices, while the deformation is modelled by the SSM (Eq. (13)). This articulated solution is finally refined by a free-form deformation with generic elasticity constraints. Further details of this iterative shape fitting procedure are provided below.

2.2.3. High-resolution single-object SSM fitting

Since the low-resolution full limb model has a unique correspondence with the detailed, high-resolution, shape models of individual segment anatomy can be

selected and local optimization of the high-resolution shape model fitting can be performed sequentially, segment by segment. Inter-penetration of surfaces at the joint sites is avoided by alternatingly masking out previously segmented neighboring segments. During fitting, the search domain is limited to the region of interest (ROI) defined by the borders of the low-resolution model. A linear gradient profile is extracted from the (vector) gradient image along each vertex normal direction. The point along this profile with the highest scalar product (exceeding a threshold value to decrease the impact of local noise) between the image gradient and the mesh normal is considered the optimal candidate point for the vertex displacement $\delta(v)$. Similar to Section 2.2.1, fitting of the SSM was again obtained by iteratively solving a set of linear regression equations mapping each model mesh vertex to its corresponding candidate image point. For each consecutive step, the positions of the model vertices obtained from the previous fitting iteration are used to define new linear profiles and new candidate corresponding points (Figure 1(C)).

2.2.4. Free form deformation

During the final fitting step, the template mesh is aligned with the target image using a free-form deformation. The deformation is defined elastically, with decreasing influence of the neighboring points (Figure 1D). Suggested displacements $\delta(v_i)$ for all template mesh points are smoothened using a local neighborhood weighing schedule:

$$d(v_i) = \sum_{j \in \aleph_i} w_{ji} * \left(\frac{1}{K} \sum_{k \in \aleph_j} \delta(v_k)\right) \tag{15}$$

with linear distance-based weights w_j defined as in Eq. (11) and K the number of points in each neighborhood. As such, a smooth estimate for mesh deformation towards the target is produced. A

Table 2. The left-hand side of the table shows the accuracy of the automatic segmentation compared with manual segmentation for each segment for $n = 10$ cases. The right-hand side reports the inter-observer variability of the manual segmentation of the respective segments.

	Automated Segmentation		Manual Segmentation	
	Accuracy ($n = 10$)		Inter-observer variability ($n = 3$)	
	Average Distance Error (mm)	Maximal Distance Error (mm)	Average Distance Error (mm)	Maximal Distance Error (mm)
Pelvis	0.75 ± 0.17	7.84 ± 2.26	0.41 ± 0.20	3.74 ± 2.68
Femur	0.65 ± 0.10	4.79 ± 2.39	0.41 ± 0.15	2.30 ± 0.98
Tibia	0.63 ± 0.11	4.07 ± 2.15	0.39 ± 0.19	1.88 ± 0.47
Patella	0.65 ± 0.15	2.02 ± 0.39	0.42 ± 0.11	1.96 ± 0.29
Fibula	0.76 ± 0.18	3.76 ± 1.17	0.61 ± 0.08	2.25 ± 0.71
Calcaneum	0.53 ± 0.16	2.90 ± 0.77	0.40 ± 0.12	1.67 ± 0.34
Talus	0.57 ± 0.12	2.97 ± 0.59	0.44 ± 0.08	2.21 ± 0.41

detailed example of the in-image segmentation process is provided in Figure 2.

2.2.5. Supervised training

Each segmented case was verified by a clinical expert. Verified segmentations were then added to the existing SSM and the model-based segmentation pipeline is progressively updated. Throughout this process, segmentation was repeated with the updated pipeline for cases that previously did not fulfill quality requirements. Four iterations with the updated SSM were finally required in total, to reach convergence in 99% of the segmentations.

2.2.6. Appearance and solids

Each bone was modelled with an outer layer, representing the periosteal surface of the bone, and an inner layer, representing the endosteal surface. During shape model construction, the periosteal and endosteal surfaces were non-rigidly registered one to the other to define dense correspondences between both layers. Image intensity at the middle of corresponding outer-inner vertices was also extracted during segmentation. The correspondence between inner and outer cortical mesh layer vertices further allows for immediate construction of tetrahedral elements (Figure 3).

2.3. Validation

Accuracy of the segmentation pipeline was established by comparing root mean squared error (RMSE) and maximum error (ME) between the automatic segmentation and the manual segmentations. Manual segmentation was performed using Mimics segmentation software (Version 17.0, Materialise NV, Heverlee, Belgium). A total of ten cases were manually segmented for this purpose. These test scans were not used to build up the shape model and unseen to the algorithm. In more detail, as manual segmentation is often prone to user misinterpretations of the boundaries of the region to be segmented, the control samples were segmented by three different individuals and the arithmetic average of the different segmentations was used to validate the accuracy and robustness of the developed segmentation pipeline.

The segmentation accuracy evolution of the SSM fitting to the in-image anatomy and the effect of the free-form optimization were evaluated as newly segmented samples were added to the SSM. Segmentation accuracy evolution was evaluated on the same 10 test cases mentioned above.

3. Results

A total of 250 CT scans were processed on a Dell Precision M6800 Laptop (Intel Core i7 -4910MQ, 16 GB RAM, 64 bit). Each scan data set consisted of an average of 1864 slices with a pixel size between 0.575 mm to 0.975 mm. Every image domain included the full lower limb anatomy from rib 12 to toes. The imaging database was constructed from living subjects receiving angio-CT scanning for vascular work-out between 2012 and 2016. CT data demonstrating metallic implants (e.g. hip and knee prosthesis) were excluded from the data base. The subjects were not exposed to additional radiation for the present study, which was approved by the local ethics committee under reference B670201111480. The dataset consisted of 148 male and 102 female subjects with an average age of 67 (±9) and 71 (±11) years, respectively.

Automatic segmentation of a full data set required on average 2 hours per case. The segmented structures included the 6 lower vertebrae, sacrum, pelvis, femur, patella, fibula, tibia, talus, calcaneum, navicular, cuboid and three cuneiform bones. A series of iterations with progressive updating of the SSM, following
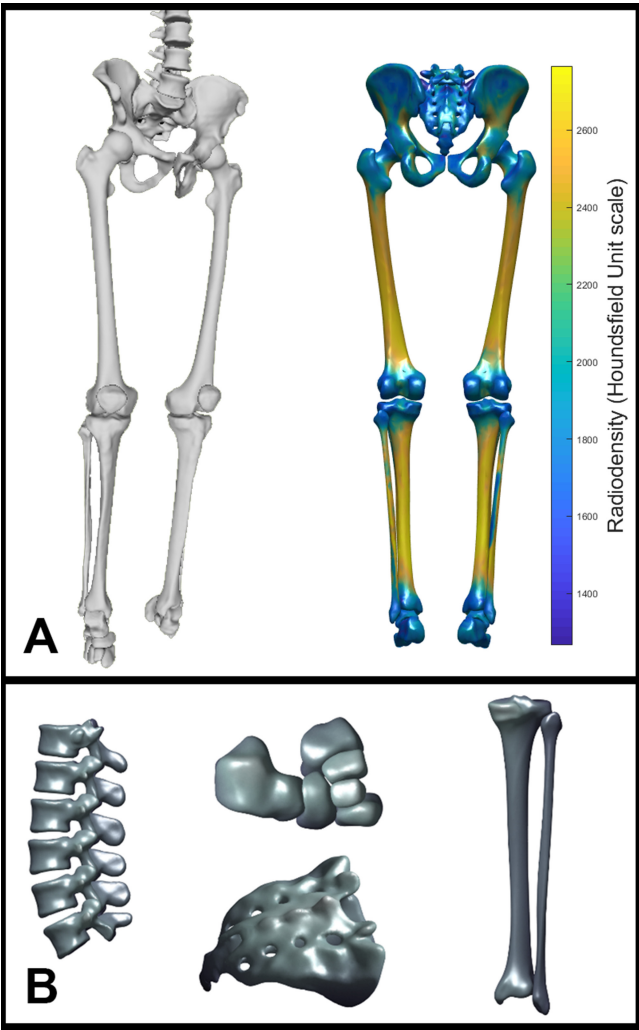
Figure 4. Average shape (left) and radiodensity (right) of the low-resolution full limb model. Radiodensity is described according to the Hounsfield scale. (A). Detail of the average shape of the high-resolution segmental parts (B).
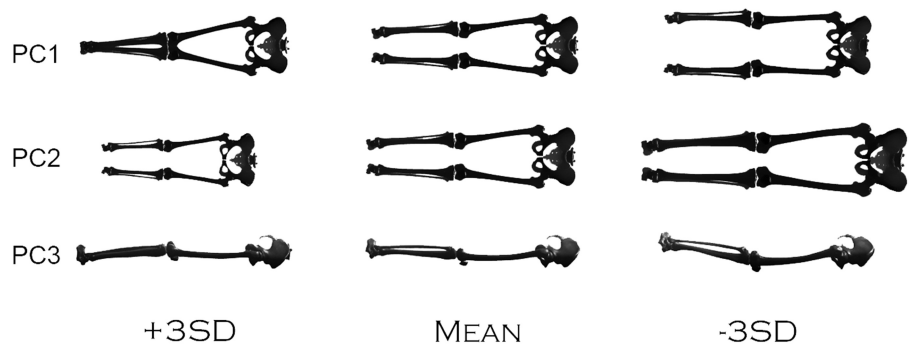

Figure 5. Representation of the first three principal modes of variation of the full limb model at the mean and at plus/minus 3 standard deviations from the mean.

quality clearance by the clinical supervisors, was performed. Finally, all scans were processed and acceptable convergence of the algorithm could be obtained in 250 cases and failed in three. Specifically, fitting of the articulated SSM failed on three cases scanned with their legs crossed.

The comparison of automatic with manual segmentations showed an RMSE ranging from 0.53 to 0.76 mm and a ME ranging from 2.0 to 7.8 mm. The highest error values were found in the pelvic bones (Table 2).

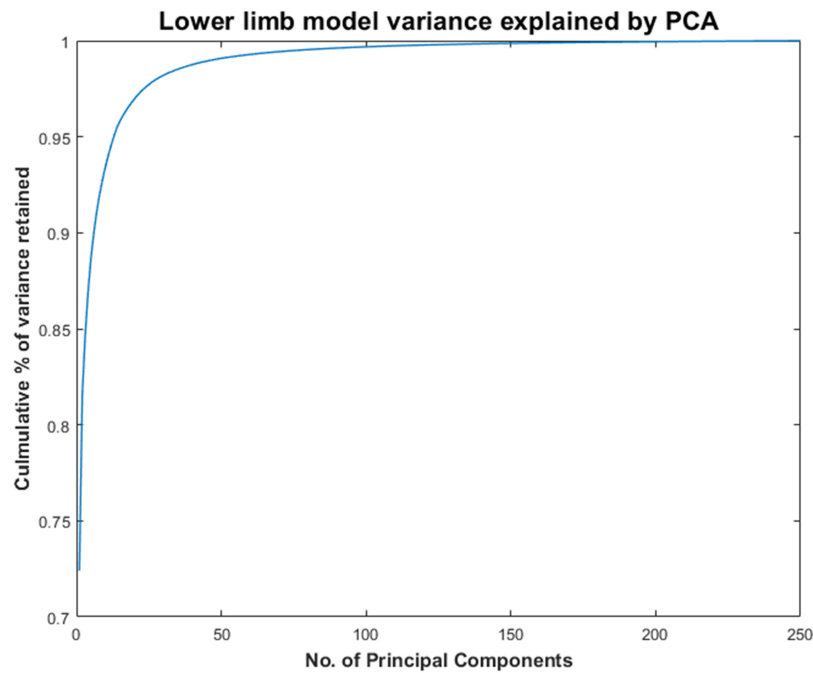Figure 4 shows the average shape and color-coded intensity models of the low-resolution model of the

Figure 6. Cumulative explained variance as a function of retained principal components following principal component analysis (PCA) to describe the articulated lower limb shape model. (X-axis: number of principal components, Y-axis: cumulative explained variance).
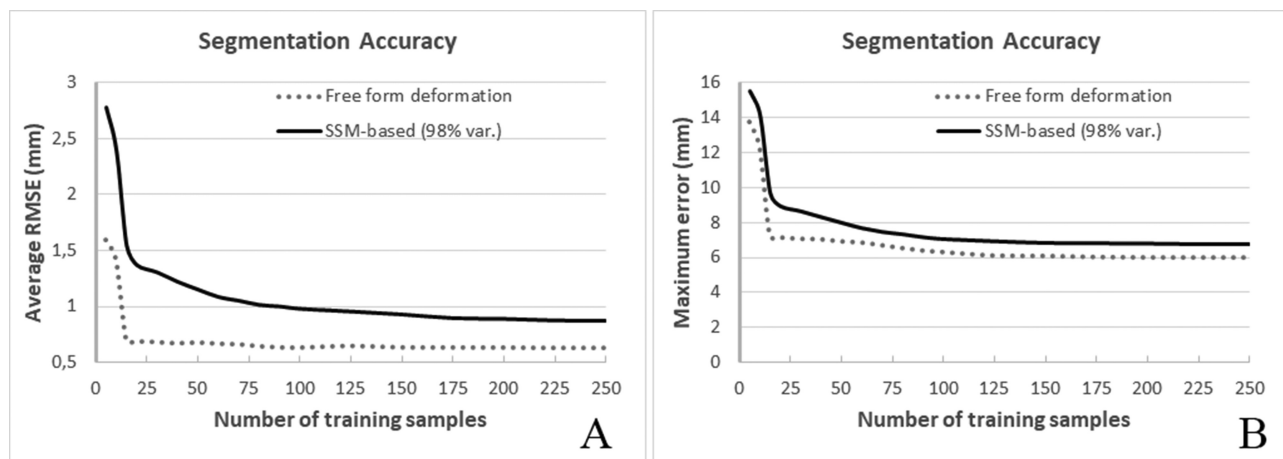

Figure 7. Segmentation accuracy evolution of the SSM-based segmentation (solid curve) and free form deformation (dotted curve) for different levels of prior knowledge expressed as amounts of training data in the SSM. Horizontal axis: nr of training samples, Vertical axis: averaged rmse (A) and maximum error (B).

full lower limb. The first 6 and 13 principal components described 90% and 95% of the cumulative variance, respectively. The first three modes of variation mainly described size and positional variation, such as crossing and bending of the legs during scanning, as show in Figure 5. The compactness of the model is illustrated by the curve of the cumulative variance explained by the principal modes (Figure 6).

With increasing number of training samples added to the articulated and single object SSMs, the accuracy of the SSM-based segmentation gradually improved. This effect was noticeable beyond 200 samples indicating the SSM did yet fully covered the population and was still improving in its generalization properties. The SSMs were implemented to describe 98% of population variance. The free-form deformation was designed to handle the remaining 2% as well as to be able to describe new anatomies to update the SSMs with. Accuracy of the free-form deformation following the SSM-based segmentation steeply inclined with

Table 3. Comparison of the current segmentation results with previously published studies. There were no automatic segmentation studies of the calcaneum and talus.

	Subjects	Segmentation method	Average distance error (mm)	Maximal Distance error (mm)
Pelvis				
Proposed method	10	SSM based	0.75 ± 0.17	7.84 ± 2.26
Seim et al. (2008)	50	SSM based	0.7 ± 0.3	16.5 ± 5
Lamecker et al. (2004)	23	SSM based	1.6 ± 0.2	14.6 ± 3.8
Kainmueller et al. (2009)	50	SSM based	0.5 ± 0.2	N/A
Chu et al. (2015)	30	Atlas and SSM based	0.4 ± 0.07	N/A
Yokota et al. (2013)	100	SSM based	1.06	N/A
Femur				
Proposed method	10	aSSM based	0.65 ± 0.10	4.79 ± 2.39
Almeida et al. (2016)	148	SSM based	1.014 ± 0.474	4.336 ± 0.861
Krcah et al. (2011)	197	Graph cut	N/A	12.5
Younes et al. (2014)	20	SSM based	1.48 ± 0.28	10.53 ± 3.19
Wu et al. (2014)	248	SSM based and Graph cut	0.82 ± 3.33	N/A
Chu et al. (2015)	60	Atlas and SSM based	0.36 ± 0.12	N/A
Kim et al. (2016)	8	Adaptive tresholding	0.36 ± 0.07	N/A
Yokota et al. (2013)	200	SSM based	0.98	N/A
Cheng et al. (2013)	110	Edge-based	0.65	N/A
Tibia				
Proposed method	10	SSM based	0.63 ± 0.11	7.84 ± 2.26
Wu et al. (2014)	248	SSM based and Graph cut	0.69 ± 1.25	N/A
Fibula				
Proposed method	10	SSM based	0.76 ± 0.18	7.84 ± 2.26
Wu et al. (2014)	248	SSM based and Graph cut	0.96 ± 4.29	N/A
Patella				
Proposed method	10	SSM based	0.65 ± 0.15	7.84 ± 2.26
Wu et al. (2014)	248	SSM based and Graph cut	0.68 ± 2.06	N/A

minimal training samples and plateaued already from 20 training samples on in terms of RMSE, however still gradually improved until 100 training samples in terms of maximum error (Figure 7).

4. Discussion

Segmentation of CT and magnetic resonance scan data has currently become the accepted standard for the development of subject-specific models. However, the segmentation process is typically manual or semi-manual and time-consuming. For instance, extracting the articular surfaces of a knee joint was reported to take approximately two days of work and full lower limb models have even been reported to require over 3 weeks of work (Boykov and Funka-Lea 2006; Henak et al. 2013).

Because of the time-consuming nature of manual segmentation, several authors have described automatic techniques as an alternative. Among these techniques, model-based approaches to segmentation are arguably the most robust methods when image data is noisy or includes artefacts (Heimann and Meinzer 2009). However, model-based segmentation requires a large number of segmented samples since segmentation accuracy has been shown to be directly related to the number of segmented samples available in the training data of the SSM (Yokota et al. 2009; Chung et al. 2011; Cheng et al. 2013). Hence, methods that

are able to process large data sets with only a limited amount of supervision, such as the proposed segmentation pipeline, are of growing interest.

An important step in the construction of corresponding meshes for use in SSM models is to define appropriate techniques for non-rigid surface registration. In this respect, numerous techniques have been proposed in the literature, all with their respective advantages and disadvantages. We have chosen to initialize the non-rigid registration process with RBF interpolation and refine with elastic deformation. RBF-based point matching and, by extension, thin plate splines, have been shown to provide robust and computationally efficient non-rigid deformation solutions by several authors (Carr et al. 2001; Chui and Rangarajan 2003; de Boer et al. 2007; Kim et al. 2016). Nevertheless, the speed of calculation comes at the cost of a reduced number of degrees of deformation freedom. For that purpose, elastic techniques have been suggested that provide solutions that do not rely on expensive interpolation (Li et al. 2008). These techniques, however, require good initial alignment and close correspondence already. Therefore, combining both techniques seemed like a logical choice to take advantage of the best of both worlds. This work indeed demonstrated that this approach leads to robust, accurate and efficient segmentation of complex anatomical structures in large data sets, such as the full lower limb in CT. The source code for

rigid and non-rigid registration has been made publicly available in the MATLAB File exchange repository (Audenaert 2013).

When compared with other methods for segmentation of bony anatomy, the proposed method performed equally well or better for the different skeletal segments (see Table 3). More specifically, it outperformed the segmentation results in the pelvis, where the maximal distance error was approximately 50% lower than previous studies (Lamecker et al. 2004; Seim et al. 2008). Femoral segmentation results reported by Almeida et al.(Almeida et al. 2016) matched well with our results. They applied a model-based segmentation technique and reached a 98% convergence rate along with a an average error of 1 mm. Wu et al. (2014) also adopted a model-based framework and included a graph-cut technique to remove overlap. The application of Wu's segmentation method on tibia, fibula and patella also corresponded very well with our results. There were no studies in the current literature on automatic segmentation of talus or calcaneum.

We could not further validate segmentation accuracy as by comparison with reported results from the literature due to the lack of availability of the imaging and segmented data used in these studies. Further, in these studies, there are large differences in CT acquisition parameters, slice thickness and voxel size which might influence the segmentation results. Further, even the gold standard of manual segmentations comes with its own limitations. Manual segmentation is also prone to errors as not only observer variability may induce misinterpretations in pixel labeling but also evaluated manual segmentation when compared with optical scanning of the bone of interest have shown average deviations of 0.55 mm. (Rathnayaka et al. 2011; Van den Broeck et al. 2014)

A second important limitation relates to the lack of real clinical scenarios such as low-resolution scans or pathological deformed joints not specifically included for evaluation. What concerns the clinical challenge of diseased joints we can add, however, that the average age of our population is close to geriatric and therefore a number of samples presented with challenging degenerative joint narrowing and structural deviations.

5. Conclusion

The current manuscript presents a technique that can be generically applied to any bony structure, and that can be initiated from a small sample of segmented bones. As such it fulfills the requirements of a segmentation pipeline that can deploy the necessary training samples for use in population wide anatomical and mechanical studies as well as for more advanced applications such as training of neural networks.

A successful bone segmentation in the majority of structures was demonstrated. The accuracy of segmentation was within the range of 0.53–0.75 mm which in line with results previously presented in the literature.

ORCID

Jan Van Houcke http://orcid.org/0000-0002-9738-7760

References

Almeida DF, Ruben RB, Folgado J, Fernandes PR, Audenaert E, Verhegghe B, De Beule M. 2016. Fully automatic segmentation of femurs with medullary canal definition in high and in low resolution CT scans. Med Eng Phys. 38(12):1474–1480.

Antony J, McGuinness K, O'Connor NE, Moran K. 2016. Quantifying Radiographic Knee Osteoarthritis Severity using Deep Convolutional Neural Networks. ICPR 2016 proceedings, 1195–1200.

Arun KS, Huang T, Fau - Blostein SD, Blostein SD. 1987. Least-squares fitting of two 3-D point sets. IEEE Trans Pattern Anal Mach Intell. 9:698–700.

Audenaert EA. 2013. nonrigidICP, MATLAB Central File Exchange. Retrieved November, 2016.

Baldwin MA, Langenderfer JE, Rullkoetter PJ, Laz PJ. 2010. Development of subject-specific and statistical shape models of the knee using an efficient segmentation and mesh-morphing approach. Comput Methods Programs Biomed. 97(3):232–240.

Besl PJ, McKay ND. 1992. A method for registration of 3-D shapes. IEEE Trans Pattern Anal Mach Intell. 14(2): 239–256. Feb

Ben Younes L, Nakajima Y, Saito T. 2014. Fully automatic segmentation of the Femur from 3D-CT images using primitive shape recognition and statistical shape models. Int J Comput Assist Radiol Surg. 9:189–196.

Bhalodia RES, Kavan L, Whitaker RT. 2018. DeepSSM: A Deep Learning Framework for Statistical Shape Modeling from Raw Images. In CW M. Reuter, H. Lombaert, B. Paniagua, M. Lüthi, B. Egger, editors. Proceedings of the ShapeMI MICCAI 2018. Berlin: Springer.

Bindernagel M. 2013. Articulated statistical shape models. Thesis. Berlin: Humboldt University.

Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N. 2008. Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans Med Imaging. 27(4):557–568.

Bourgeat P, Fripp J, Stanwell P, Ramadan S, Ourselin S. 2007. MR image segmentation of the knee bone using phase information. Med Image Anal. 11(4):325–335. Aug

Boykov Y, Funka-Lea G. 2006. Graph cuts and efficient N-D image segmentation. Int J Comput Vision. 70(2): 109–131.

Carr JC, Beatson RK, Cherrie JB, Mitchell TJ, Fright WR, McCallum BC, Evans TR. 2001. Reconstruction and representation of 3D objects with Radial Basis Functions. Siggraph 2001 Conference Proceedings, 67–76.

Cheng YZ, Zhou SJ, Wang YD, Guo CY, Bai J, Tamura S. 2013. Automatic segmentation technique for acetabulum and femoral head in CT images. Pattern Recogn. 46(11): 2969–2984. Nov

Chu C, Chen C, Liu L, Zheng G. 2015. FACTS: Fully Automatic CT Segmentation of a Hip Joint. Ann Biomed Eng. 43(5):1247–1259.

Chui HL, Rangarajan A. 2003. A new point matching algorithm for non-rigid registration. Comput Vis Image Underst. 89(2–3):114–141.

Chung F, Schmid J, Magnenat-Thalmann N, Delingette H. 2011. Comparison of statistical models performance in case of segmentation using a small amount of training datasets. Vis Comput. 27(2):141–151.

Claes P. 2007. A robust statistical surface registration framework using implicit function representations: application in craniofacial reconstruction. Thesis. Leuven: Katholieke Universiteit.

Cootes TF, Hill A, Taylor CJ, Haslam J. 1994. Use of active shape models for locating structure in medical images. Image Vis Comput. 12(6):355–365.

Cootes TF, Taylor CJ, Cooper DH, Graham J. 1995. Active shape models - their training and application. Comput Vis Image Underst. 61(1):38–59. Jan

de Boer A, van der Schoot MS, Bijl H. 2007. Mesh deformation based on radial basis function interpolation. Comput Struct. 85:784–795. Jun

Esat II, Bahai H. 2000. Surface alignment based on the moment of inertia and improved least-squares methods. Proc Inst Mech Eng B J Eng Manuf. 214(7):547–554.

Gassman EE, Powell SM, Kallemeyn NA, DeVries NA, Shivanna KH, Magnotta VA, Ramme AJ, Adams BD, Grosland NM. 2008. Automated bony region identification using artificial neural networks: reliability and validation measurements. Skeletal Radiol. 37(4):313–319. Apr

Heimann T, Meinzer HP. 2009. Statistical shape models for 3D medical image segmentation: A review. Med Image Anal. 13(4):543–563.

Henak CR, Anderson AE, Weiss JA. 2013. Subject-specific analysis of joint contact mechanics: application to the study of osteoarthritis and surgical planning. J Biomech Eng -Trans ASME. 135:021003.

Kainmueller D, Lamecker H, Zachow S, Hege HC. 2008. Coupling deformable models for multi-object segmentation. Proc. Biomed. Simul. 5104:69–78.

Kainmueller D, Lamecker H, Zachow S, Hege HC. 2009. An articulated statistical shape model for accurate hip joint segmentation. Conf Proc IEEE Eng Med Biol Soc. 2009:6345+.

Kim Y, Na YH, Xing L, Lee R, Park S. 2016. Automatic deformable surface registration for medical applications by radial basis function-based robust point-matching. Comput Biol Med. 77:173–181.

Krcah M, Szekely G, Blanc R, Ieee. 2011. Fully Automatic and Fast Segmentation of the Femur Bone from 3D-CT Images with no Shape Prior. In: 2011 8th Ieee International Symposium on Biomedical Imaging: From Nano to Macro. New York: Ieee. p. 2087–2090.

Lamecker H, Seebass M, Hege HC, Deuflhard P. 2004. A 3D statistical shape model of the pelvic bone for segmentation. In: Medical Imaging 2004: Image Processing, Pts 1-3. Bellingham: Spie-Int Soc Optical Engineering; p. 1341–1351.

Li H, Sumner RW, Pauly M. 2008. Global correspondence optimization for non-rigid registration of depth scans. Comput Graph Forum. 27(5):1421–1430.

Li K, Wu XD, Chen DZ, Sonka M. 2006. Optimal surface segmentation in volumetric images – a graph-theoretic approach. IEEE T Pattern Anal. 28:119–134.

Liu F, Zhou ZY, Jang H, Samsonov A, Zhao GY, Kijowski R. 2018. Deep convolutional neural network and 3D deformable approach for tissue segmentation in musculoskeletal magnetic resonance imaging. Magn Reson Med. 79(4):2379–2391. Apr

Otsu N. 1979. Threshold Selection Method from Gray-Level Histograms. IEEE Trans Syst, Man, Cybern. 9(1):62–66.

Rathnayaka K, Sahama T, Schuetz MA, Schmutz B. 2011. Effects of CT image segmentation methods on the accuracy of long bone 3D reconstructions. Med Eng Phys. 33(2):226–233.

Ronneberger O, Fischer P, Brox T. 2015. U-Net: convolutional networks for biomedical image segmentation. Med Image Comput Comput Assist Interv. 9351:234–241.

Seim H, Kainmueller D, Heller M, Lamecker H, Zachow S, Hege HC. 2008. Automatic segmentation of the pelvic bones from CT data based on a statistical shape model. Eurographics Workshop on Visual Computing for Biomedicine.

Shim H, Chang S, Tao C, Wang JH, Kwoh CK, Bae KT. 2009. Knee cartilage: efficient and reproducible

segmentation on high-spatial-resolution MR images with the semiautomated graph-cut algorithm method. Radiology. 251(2):548–556.

Tikhonov AN, Arsenin VY. 1977. Solution of Ill-posed problems. Washington: Winston & Sons.

Van den Broeck J, Vereecke E, Wirix-Speetjens R, Sloten JV. 2014. Segmentation accuracy of long bones. Med Eng Phys. 36(7):949–953. Jul

Wu DJ, Sofka M, Birkbeck N, Zhou SK. 2014. Segmentation of Multiple Knee Bones from CT for Orthopedic Knee Surgery Planning. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2014, Pt I, 372–380.

Yin Y, Zhang XM, Williams R, Wu XD, Anderson DD, Sonka M. 2010. LOGISMOS-layered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint. IEEE T Med Imaging. 29: 2023–2037. Dec

Yokota F, Okada T, Takao M, Sugano N, Tada Y, Sato Y. 2009. Automated segmentation of the femur and pelvis from 3D CT data of diseased hip using hierarchical statistical shape model of joint structure. Medical image computing and computer-assisted intervention – Miccai 2009, Pt Ii, Proceedings. 5762:811.

Yokota F, Okada T, Takao M, Sugano N, Tada Y, Tomiyama N, Sato Y. 2013. Automated CT segmentation of diseased hip using hierarchical and conditional statistical shape models. Medical image computing and computer-assisted intervention - Miccai 2013, Pt Ii. 8150: 190–197.