Medical & Biological Engineering & Computing
https://doi.org/10.1007/s11517-023-02882-3

ORIGINAL ARTICLE

3D biplanar reconstruction of lower limbs using nonlinear statistical models

Dac Cong Tai Nguyen[1,2] · Said Benameur[2] · Max Mignotte[1] · Frédéric Lavoie[2,3]

Received: 25 November 2022 / Accepted: 20 June 2023

Abstract
Three-dimensional (3D) reconstruction of lower limbs is of great interest in surgical planning, computer assisted surgery, and for biomechanical applications. The use of 3D imaging modalities such as computed tomography (CT) scan and magnetic resonance imaging (MRI) has limitations such as high radiation and expense. Therefore, three-dimensional reconstruction methods from biplanar X-ray images represent an attractive alternative. In this paper, we present a new unsupervised 3D reconstruction method for the patella, talus, and pelvis using calibrated biplanar (45- and 135-degree oblique) radiographic images and a prior information on the geometric/anatomical structure of these complex bones. A multidimensional scaling (MDS)-based nonlinear dimensionality reduction algorithm is applied to exploit this prior geometric/anatomical information. It represents relevant deformations existing in the training set. Our method is based on a hybrid-likelihood using regions and contours. The edge-based notion represents the relation between the external contours of the bone projections and an edge potential field estimated on the radiographic images. Region-based notion is the non-overlapping ratio between segmented and projected bone regions of interest (RoIs). Our automatic 3D reconstruction model entails stochastically minimizing an energy function allowing an estimation of deformation parameters of the bone shape. This 3D reconstruction method has been successfully tested on 13 biplanar radiographic image pairs, yielding very promising results.

Keywords 3D reconstruction · X-ray images · Biplanar radiographies · Nonlinear statistical models · Medical imaging

1 Introduction

Three-dimensional (3D) reconstruction of bones is important for diagnosis and treatment of pathologies of the bony structure of the human body, especially for precise implant selection. Precise personalized 3D models have great potential as regards accuracy and reliability in orthopaedics. X-ray images are often used due to their wide availability, lower price, and low levels of ionizing radiation compared to computed tomography (CT) scan and magnetic resonance imaging (MRI). However, since X-ray images provide only 2D information, some prior knowledge must be incorporated in order to model the third dimension.

Previous approaches [1–6] use *a priori* knowledge of the geometric structure of the object to be reconstructed. Nevertheless, in these methods, the geometric *a priori* constraint does not model the set of admissible deformations of the anatomical structure to be reconstructed. Consequently, the 3D shape estimation does not reliably correspond to the reality.

Methods using statistical *a priori* knowledge of the geometric shapes of the objects of interest lead to better constrain the reconstruction problem [7–11]. These methods use statistical shape models or statistical shape and intensity models for reconstructing bones from X-ray images. However, the optimization of the deformable model parameters can be slow and requires a good initialization point to avoid local maxima [12, 13]. [14, 15] used the contours of anatomical shapes detected in the two radiographic images for the 3D reconstruction of the anatomical object. However, segmentation or edge detection algorithms are very sensitive to noise, and this often results in noisy contours, which are difficult to manage in radiographic images.

[16] used generative adversarial network (GAN) to reconstruct 3D spine structure from simulated biplanar X-ray images. Unfortunately, it was a basic prototype model with constrained training portfolios. In addition, 2D simulated X-ray images derived from 3D structures lost some exact

✉ Dac Cong Tai Nguyen
dac.cong.tai.nguyen@umontreal.ca

Extended author information available on the last page of the article

Published online: 28 July 2023

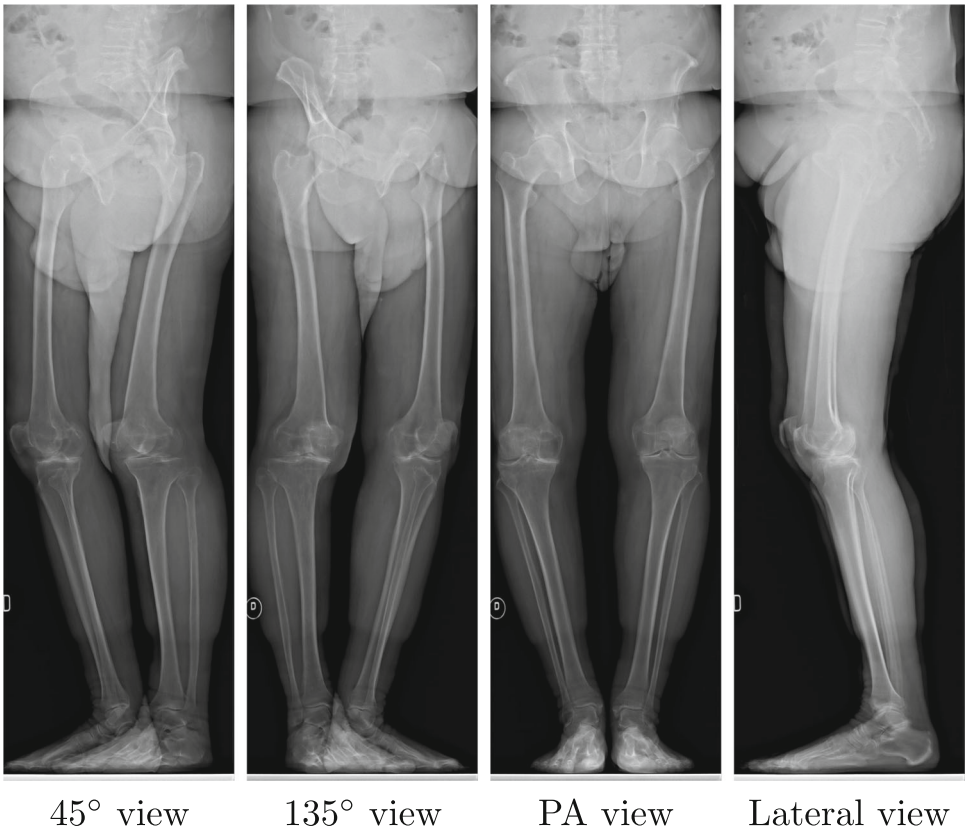


Fig. 1 Example of biplanar images. The bone structures in oblique images are less clear than in a PA image but much clearer than in a LAT image

information. [17] used convolutional neural network (CNN) for 3D reconstruction of bones from two biplanar X-ray images. However, due to the transparent nature of X-ray images, matching surface points among multi-views for dense reconstruction is extremely challenging compared to the standard multi-view setting. Due to the difficulty of representing a dimensional enlargement in multi-view settings with standard differentiable layers, 3D reconstruction using a deep learning approach remains a challenging task. [18] presented a survey on deep learning in medical image reconstruction. Deep learning methods require considerable computing power, are expensive to set up, make decisions that are difficult or not at all understandable, and require a large database. Maken et al. [19] give a comparative review on 3D reconstruction from X-ray images for clinical purposes.

The 3D reconstruction approach for bones that we propose uses calibrated 45- and 135-degree oblique radiographic images[1] and a general *a priori* knowledge of the geometric/anatomical structure of each bone. This knowledge comes from a considerable learning base of surface models. This *a priori* knowledge is obtained by applying the variant of multidimensional scaling (MDS) algorithm [20, 21] and by choosing as the metric the most adequate distance to our learning base. MDS allows finding a small-dimensional subspace which preserves the chosen metric in the original space. This step of nonlinear dimensionality reduction allows us to define a mesh of deformation in a reduced dimensional space in which each point of this mesh will correspond to a surface model of our learning base. As for the principal component analysis (PCA) in [14], it allows us to define a concise 3D parametric reconstruction model in which the set of statistically admissible (in our case, possibly nonlinear) deformations, confronted with our anatomical learning database, will be summarized by the values of a reduced vector of parameters.

From this mesh of nonlinear deformations which summarizes all the statistically admissible surface structures, the proposed reconstruction method consists in adjusting the projections (45-degree and 135-degree oblique) of a surface model of bone (contained in this deformation mesh and by interpolation) with the contours and the regions of segmented images, containing the corresponding bone. This approach is based on a likelihood using *a prior* detection of the contours (oblique projections of the 3D model) as in [14] but also on a global constraint using the notion of specificity. This property is based on the fact that labels (of classes), given by a prior textural over-segmentation of the image, inside and outside the shape, are distinct (or specific in a neighborhood

[1] Due to the physical link between the detector source assemblies, the position in space of the sensors and X-ray sources are well known: the radiographic environment is therefore pre-calibrated.

of the object). Thanks to this likelihood, using both edge and region information, our Bayesian reconstruction model has the property of being particularly robust against noise. Finally, the 3D reconstruction problem is thus seen as a simple problem of estimating the deformation parameters of this 3D surface model or, equivalently, as a problem of minimizing a cost function. This minimization is efficiently carried out by a stochastic algorithm.

The remainder of this paper is divided into the following sections: Sect. 2 describes our approach with the pre-processing step made on each X-ray image to enhance the contours and extract the regions, the statistical deformable model, energy function, and the optimization algorithm. We show experimental results in Sect. 3, discuss in Sect. 4, and conclude in Sect. 5.

2 Proposed model

Our 3D reconstruction approach uses two calibrated biplanar oblique 45-degree and 135-degree radiographic images and *a prior* global knowledge of the geometric/anatomical structure of each bone. We reason that the oblique orthogonal radiologic projections will allow a better identification of the various components of the bone.

Note that most biplanar 3D reconstruction methods use postero-anterior (PA) and lateral (LAT) images. The human bone structures in the PA image are clearly identified, usually non-overlapping and distinguished from dense soft tissue (cartilage, meniscus, and fascia). But in the LAT image, due to the overlapping of bones and dense soft tissue, much less information about bone structure can be extracted. In contrast, the bone structures in both oblique images appear less clear than in a PA image but much clearer than in a LAT image. Bone structure information can be equally extracted from each of two oblique images (see Fig. 1).

A robust approach consists of a segmentation method which makes it possible to take into account both the notions of contour and region (detection of homogeneous zones). The combination of segmentation by region and contour contributes to a better appreciation of the bones in the radiographic images and, consequently, to a better segmentation [22] thanks to the complementary nature of these two types

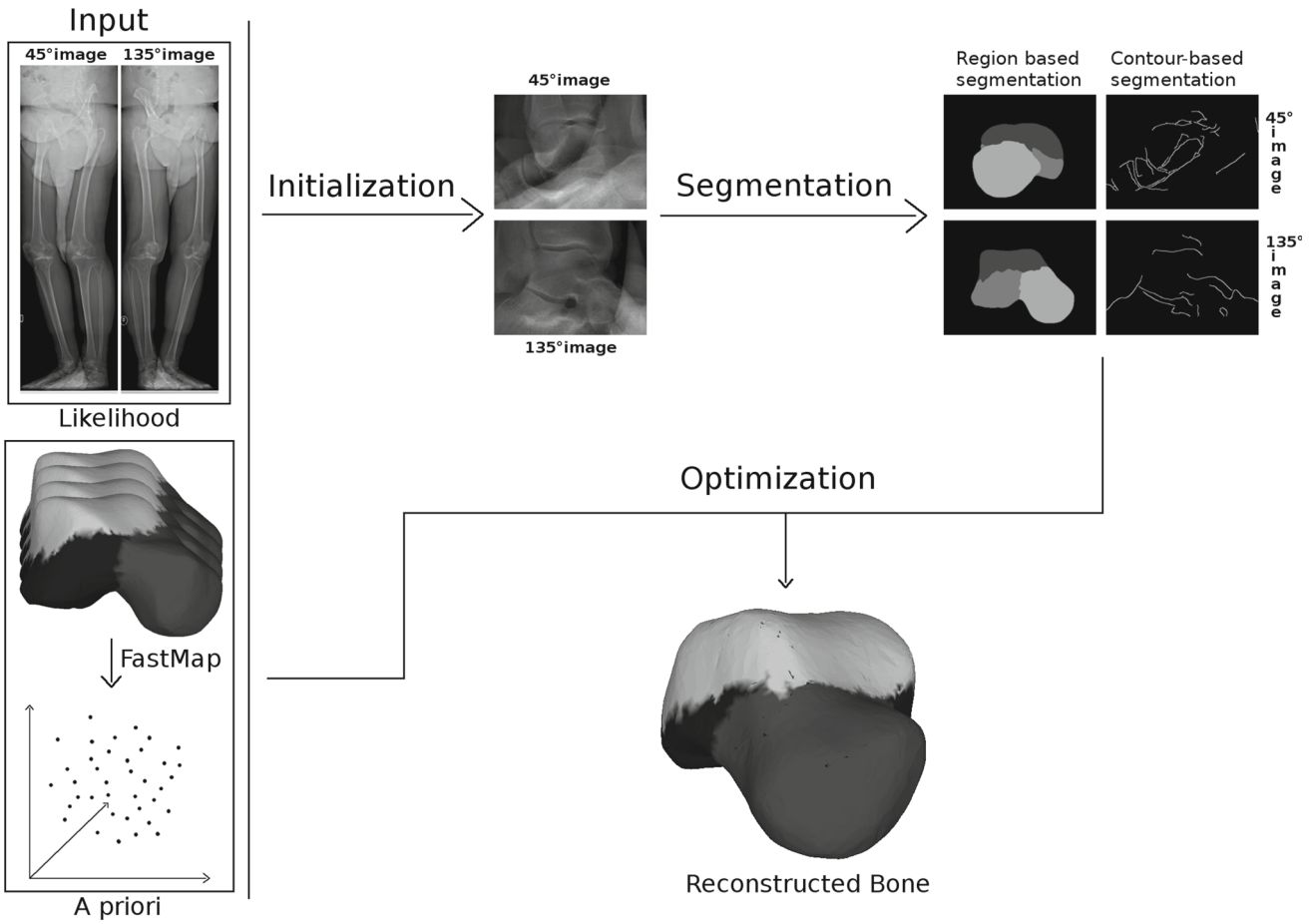


Fig. 2 The pipeline of the proposed method

of information [23]. Thus, segmentation by region and contour cooperation/combination can be thought of as mutual aid between these two concepts in order to improve the end result [24, 25].

The prior information is obtained by applying FastMap [26], a variant of the MDS algorithm, on the training set (in which each 3D shape is represented by a fixed-length vector, representing the coordinates of its different points). This algorithm uses an adequate distance metric to find an adequately low-dimensional space that preserves the chosen metric in the original space. This nonlinear dimensionality reduction step allows us to define a mesh of deformations in a reduced dimensionality space in which each point corresponds to an element of our training set. In our application, this mesh also defines a concise nonparametric model of the possible 3D deformations in which the admissible statistical deformations, extracted from the training set, are given by the set of values of the different reduced parameter vectors.

From this triangular mesh of nonlinear deformations which summarizes all the *a priori* admissible statistical structures, the objective of the proposed approach is to (non-parametrically) sample a population of possible (i.e., statistically admissible) deformed 3D shapes to be reconstructed. To this end, a sampling strategy is performed from this density mesh by using the interpolation of the nearest (and not dimensionally reduced) k closest elements of this mesh. An individual fitness value for each candidate shape, sampled from this population, is calculated from the projected contours and regions of interest (RoIs) on the biplanar oblique radiographic images with the corresponding bone in the preprocessed images (and after the deformation shape is adjusted by a deterministic optimization algorithm). Finally, the set of fitness values for this population is used to effectively and iteratively guide a stochastic optimization process to increasingly promising solutions until the optimal reconstructed shape is found (see Fig. 2).

2.1 Training set

We have 3D databases for 654 patella, 380 talus, and 39 pelvis. These databases were obtained from the semi-automated segmentation of CT images [27]. This semi-automated segmentation method was based on the propagation of the contours in adjacent slices. An initial contour was outlined by manually selecting few high curvature points and using the Fourier interpolation method to complete the contour. Then, a deformable image registration method was applied to map the contour to adjacent slices. Next, for each bone, the surface models were registered by a rigid registration approach called Iterative Closest Point (ICP) [28] that iteratively minimizes the distance metric in the least squares error. Finally, we calculated the corresponding points between the models of a bone using descriptors, namely, a set of features, assigned to a point and describing the local geometry around it [29]. For each bone, all the models had the same $\mathcal{N}$ number of 3D points. These points were indexed in such a way that the detection of RoIs was easy for all bones.

Figure 3 represents the regions involved in the segmentation process of the patella, the talus, and the pelvis, respectively. These RoIs were defined by an orthopedic surgeon.

2.2 Image preprocessing

Contour detection

Preprocessing is performed on each image to enhance the visibility of the component contours, and this constitutes the most important and reliable low-level visual cue in each radiographic image. First, a histogram equalization technique increases the global contrast. Second, a median filter and non-local means denoising algorithm [30] are used to remove

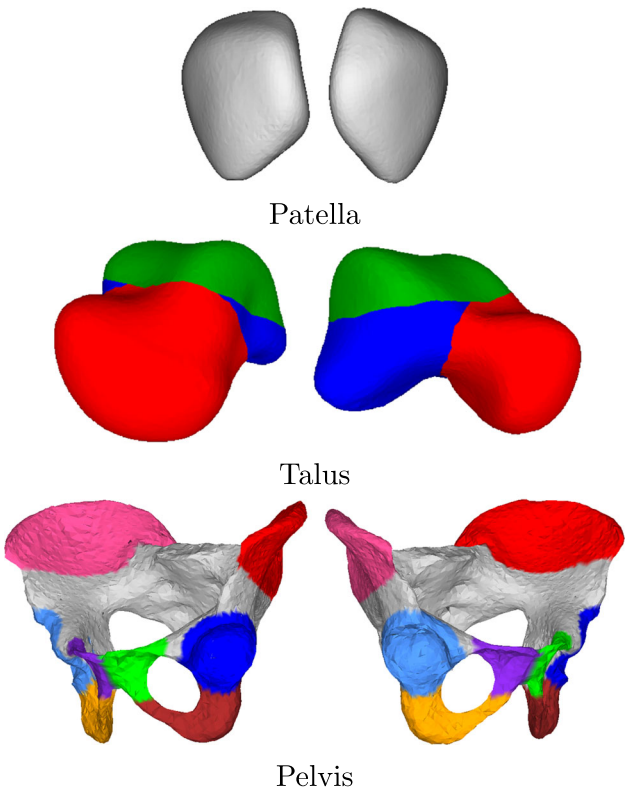


Fig. 3 Example of 3D bone regions. 45-degree view is on the left, and 135-degree view is on the right of each bone. The patella has 1 region, the talus is divided into 3 regions (green: dome, blue: body, red: head), and the pelvis is divided into 9 regions (pink and red: right and left iliac spine, cyan and blue: right and left acetabulum, purple and green: right and left pubis, orange and brown: right and left ischium, grey: rest of pelvis)

the noise from the images. The non-local means denoising method replaces a pixel with a weighted average of pixels having a similar neighborhood. More precisely, for each pixel, it first searches in a large search window (centered on the pixel to be denoised) for all the neighborhoods of pixels that most closely resemble (with a least squares (LSQ) similarity measure) the neighborhood of the pixel to be filtered. Then, a weighted average (based on the previous LSQ similarity measure) of all these central pixels (of all these neighborhoods) allows to estimate the denoised greyscale value of the pixel. Finally, the edges are detected by using a Canny edge filter [31] (see Fig. 4).

Region detection

The next part of the preprocessing step is the region extraction. By using the superpixel and multi-atlas-based algorithm [22], the input images are segmented into *m* RoIs. This algorithm registers each image in the training set to the input image using a contour-based registration technique. A superpixel map is estimated from the first half of the optimally registered images in the training set by simply taking the intersection of all the regions existing in this selected subset. Each superpixel represents the (local) variability of bone regions. This map is adaptive to the input image and takes into account all the nonlinear and local variability of bone regions existing in the selected subset. Then, a pruning step is achieved by finding the set of connected superpixels which maximize the contour-based similarity between the outer contour of this superpixel map and the edge map of the input image. Finally, a label propagation step based on the entropy concept is used for refining the resulting segmentation map into the most likely internal regions to the final consensus segmentation (see Fig. 5).

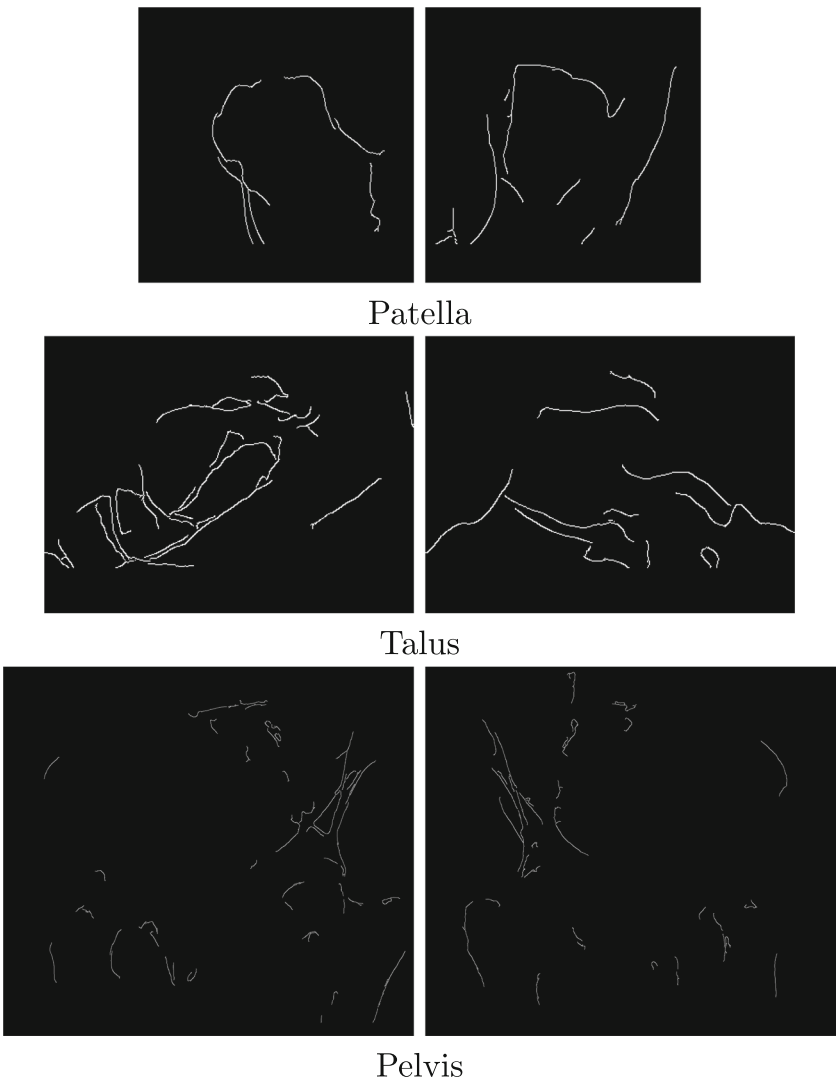


Fig. 4 Example of bone contour detection using our Canny filter-based approach. 45-degree view is on the left, and 135-degree view is on the right of each bone

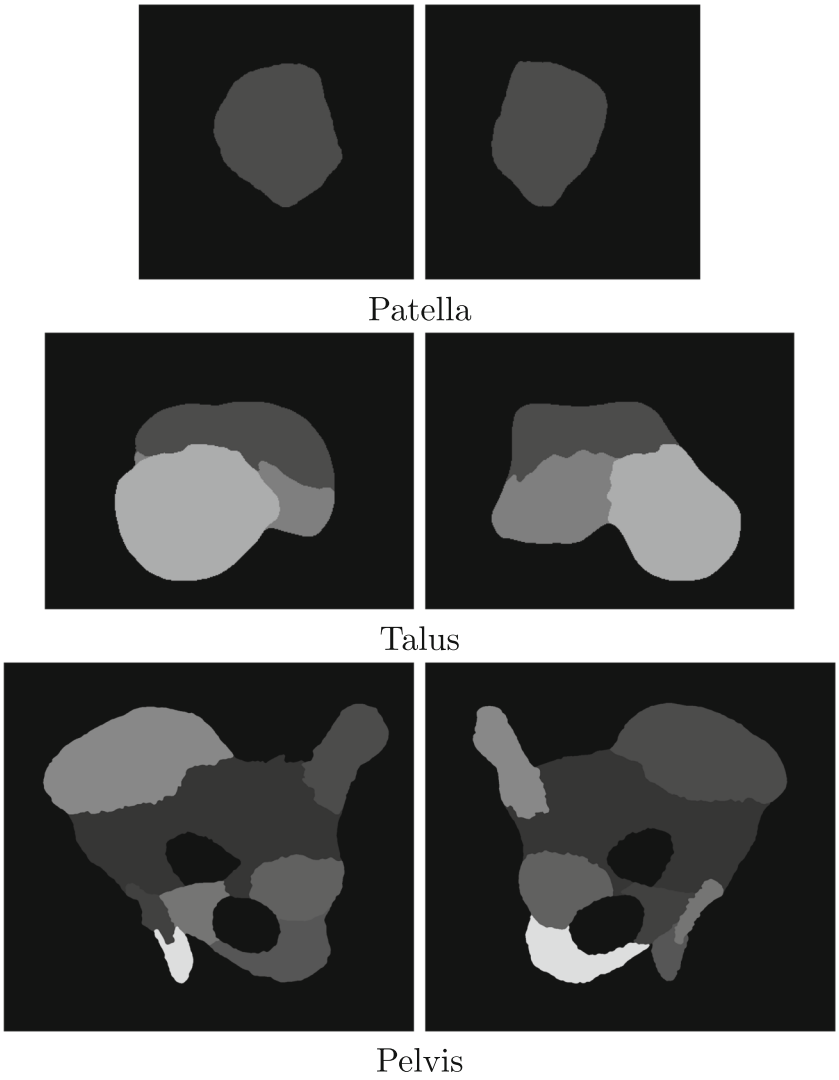


Fig. 5 Example of bone region detection using the superpixel and multi-atlas-based algorithm. The 45-degree view is on the left, and 135-degree view is on the right of each bone

2.3 Statistical deformable model

This MDS-based algorithm has been successfully used in many imaging applications, such as 3D object recognition [32], human action recognition [33, 34], image segmentation [35, 36], and image change detection [37]. In our study, each object in the training set is considered to be a point in $\mathcal{N}$-dimensional space. These points are then projected on $\mathcal{M}$ mutually orthogonal axes. The goal is to calculate the training set in an $\mathcal{M}$-dimensional space from the distance matrix previously computed in the original space ($\mathcal{M} << \mathcal{N}$).

FastMap algorithm

is based on the projections of the objects on a selected line. This line is created by connecting two furthest pivot objects O_a and O_b in the training set (see Algorithm 1). Then, the projections of the objects O_i are computed by applying the cosine law defined as

$$\mathcal{D}^2(O_b, O_i) = \mathcal{D}^2(O_a, O_i) + \mathcal{D}^2(O_a, O_b) - 2x_i\mathcal{D}(O_a, O_b) \quad (1)$$

where $\mathcal{D}(O_i, O_j)$ (for $i, j = 1, ..., \mathcal{N}$) is the distance between O_i et O_j. From the Pythagorean theorem, Eq. 1 can be solved for x_i, the first coordinate of object O_i:

$$x_i = \frac{\mathcal{D}^2(O_a, O_i) + \mathcal{D}^2(O_a, O_b) - \mathcal{D}^2(O_b, O_i)}{2\mathcal{D}(O_a, O_b)} \quad (2)$$

For $\mathcal{M} > 1$, the objects are projected on a hyper-plane $\mathcal{H}$ which is perpendicular to the line O_aO_b. A new distance $\mathcal{D}'$ between two projections O_i' et O_j' on $\mathcal{H}$ is deduced from the original distance $\mathcal{D}$ as follows:

$$\mathcal{D}'^2(O_i'O_j') = \mathcal{D}^2(O_iO_j) - (x_i - x_j)^2 \quad (3)$$

Then, Eq. 2 (with new distance function $\mathcal{D}'$, and new pivot objects) is applied to obtain the next coordinate of objects (i.e., dimension $\mathcal{M} > 1$).

The efficiency of FastMap, in terms of preserving information in reduced dimensionality, can be evaluated by calculating a correlation metric [38]. This metric is the correlation of Euclidean distance of

Algorithm 1 Heuristic to choosing two distant objects

begin

- Chose arbitrarily an object and declare it to be the second pivot object O_b
- $O_a \leftarrow$ the farthest object from O_b according to the distance function $\mathcal{D}$
- $O_b \leftarrow$ the farthest object from O_a according to the distance function $\mathcal{D}$
- return O_a and O_b as the desired pair of object

end

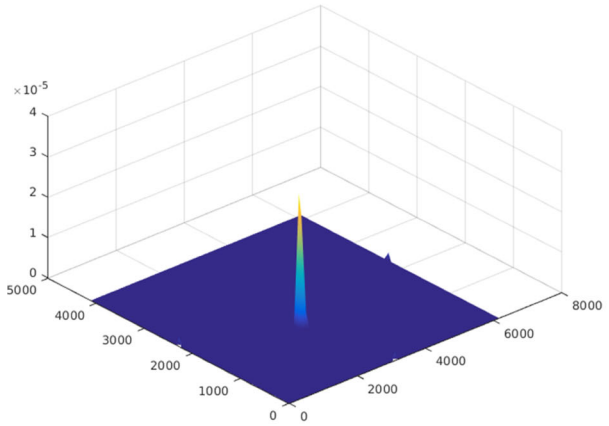
each object pair in the original space and their corresponding pair in reduced dimensionality space:

$$\text{correlation}(X, Y) = \frac{\text{covariance}(X, Y)}{\sigma_X \sigma_Y} = \frac{\frac{X^t Y}{|X|} - \overline{X}\,\overline{Y}}{\sigma_X \sigma_Y} \tag{4}$$

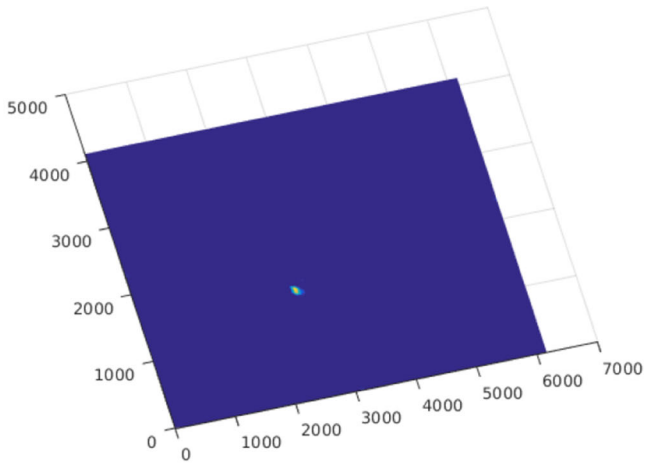
where $X and Y$ are respectively the vector of distance in the original space and in the reduced dimensionality space. X^t, $|X|$, $\overline{X}$, and σ_X are the transpose, the cardinal, the mean, and the standard deviation of X, respectively.

This correlation specifically quantifies the degree of dependence between X and Y and shows how FastMap is able to give a cartography in the reduced dimensionality space in which each point is placed such as the inter-point distances (in the original space) are preserved as much as possible [39]. A perfect correlation of 1 gives a perfect (positive) linear correlation or relation between the high dimensional data and the low dimensional data (i.e., no loss of information), and a correlation of 0 means a total loss of information (e.g., a correlation of 0.80 means that the FastMap technique succeeds in keeping 80% identical in terms of distance of the object pairs between the two [original and reduced] spaces).

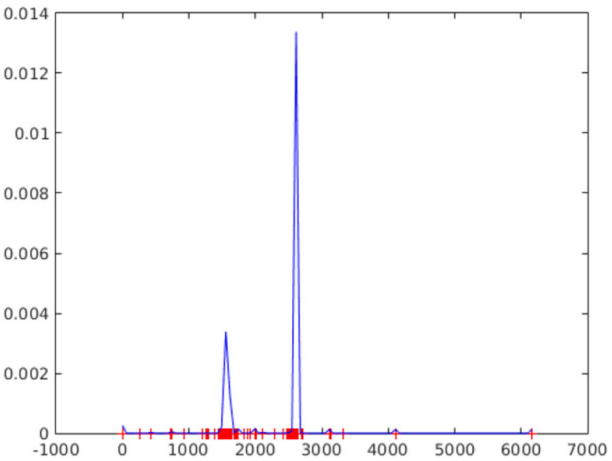
Figures 6, 7, and 8 show the cartographies or density mappings of patella, talus, and pelvis shapes at dimension $\mathcal{M} = 1$ and at $\mathcal{M} = 2$, respectively, and showing for a complex shape like the pelvis, several modes of deformation.



(a) $\mathcal{M} = 2$, profile view



(b) $\mathcal{M} = 2$, top view



(c) $\mathcal{M} = 1$

Fig. 6 Distribution of patella in low dimensionality space

Table 1 shows the correlation metric of patella, talus, and pelvis at different reduced dimensions.

For each possible (or statistically admissible) candidate 3D deformed shape, which will be sampled from the preliminary estimated shape mappings (in reduced dimensionality), the proposed model will find the k nearest neighbors of this sample in the low-dimensional data and then will use these k corresponding objects in the high-dimensional space to generate, by interpolation (by applying the inverse distance weighting (IDW) function), a new instance of the S bone shape. The weight w is an inverse distance of a point to its nearest neighbors and is computed in the reduced dimensionality space. The new bone shape is calculated in the original space.

$$S(z) = \frac{\sum_i^k w_i * Z_i}{\sum_i^k w_i} \tag{5}$$

$$w_i = \frac{1}{dist(z, z_i)^{\varrho}} \tag{6}$$

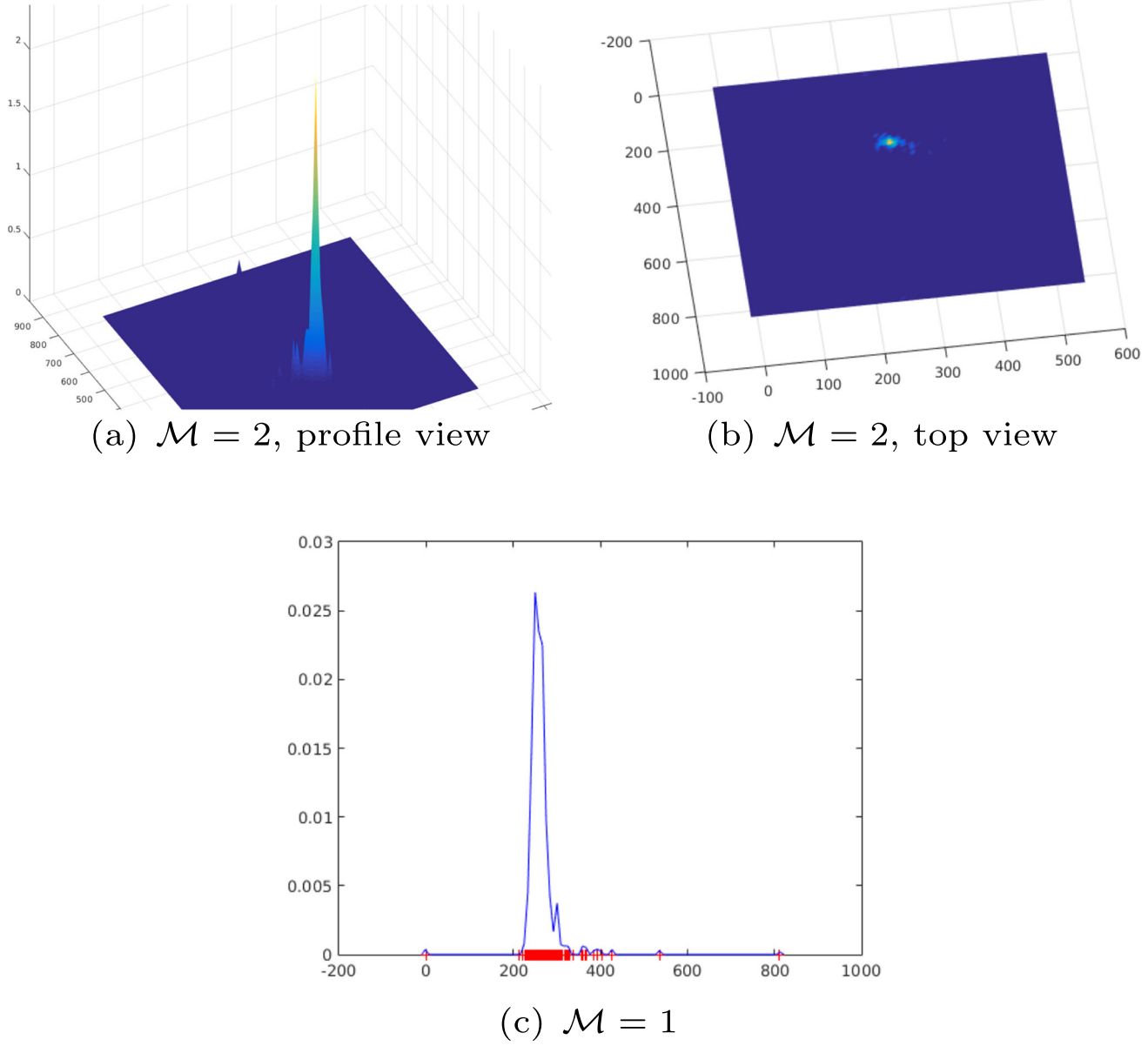


Fig. 7 Distribution of talus in low dimensionality space

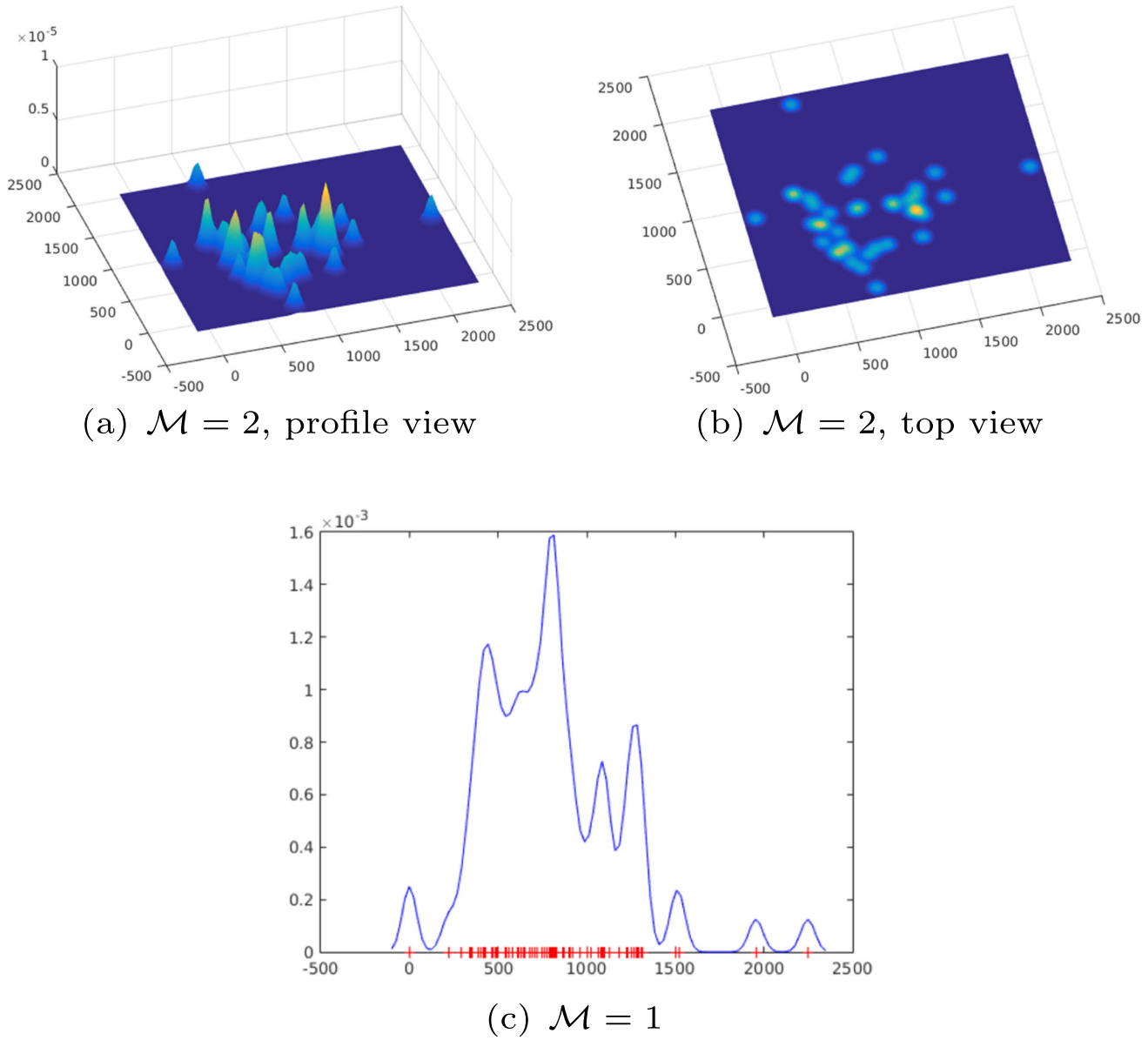


Fig. 8 Distribution of pelvis in low dimensionality space

where z is a point in the reduced dimensionality space, $dist$ is the distance between 2 points, $\varrho \geq 0$ is the power parameter, and Z is an object in the (original) high dimensional space.

2.4 Energy function term

2.4.1 Likelihood energy term

Edge potential field-based similarity

An edge potential field-based similarity measure evaluates the concordance or the similarity between the external contours of the bone silhouette projections on the two biplanar X-ray images and an edge potential field, calculated from the previously detected contours. This edge potential field attracts the bone contours and aligns them on the edge of the input image, by giving (concretely) a similarity measure proportionally greater as the projected contours coincide well with the edges existing in the images. In addition, a directional component is added to complete the measurement of the correspondence between the projected contours of the human bone and the edges in the two views. This measure is computed on the preprocessed image and is defined as [40]:

$$\xi(S) = \frac{1}{n}\sum_{n}(\Phi(x, y)|\cos(\alpha(x, y))|) \tag{7}$$

$$\Phi(x, y) = \exp(-\rho\sqrt{\delta_x^2 + \delta_y^2}) \tag{8}$$

where (δ_x, δ_y) is the displacement to the nearest edge point in the image, ρ is a smoothing factor which controls the degree of smoothness of the potential field Φ, $\alpha(x, y)$ is the angle between the tangent

Algorithm 2 FastMap

$\mathcal{O}$	set of objects in original dimensional space
$\mathcal{M}$	reduced dimensionality
$\mathcal{Y}$	set of objects in reduced dimensional space

1. Initialization
$t \leftarrow 0$

2. Algorithm Fastmap($\mathcal{M}$, $\mathcal{D}()$, $\mathcal{O}$)
if $\mathcal{M} \leq 0$ **then**
  return
end if
$t \leftarrow t + 1$
Choose pivots $0_a, 0_b$ by Algorithm 1
if $\mathcal{D}(0_a, 0_b) == 0$ **then**
  $\mathcal{Y}[i, t] = 0$ for every i
  return
end if
for each $O_i \in \mathcal{O}$ **do**
  Compute x_i using Eq. 2
  $\mathcal{Y}[i, t] = x_i$
end for
Call Fastmap($\mathcal{M} - 1$, $\mathcal{D}'()$, $\mathcal{O}$)

direction of the projected external contours at (x, y) and the tangent of the nearest edge, and n is the number of pixels on projected external contours of the human bone S (see Fig. 9).

Non-overlapping ratio-based similarity

The non-overlapping ratio evaluates the matching between regions of the segmented image and the bone silhouette projection on biplanar X-ray images. This ratio is simply the number of pixels of segmented and projected regions that are missing from their intersection on the number of pixels of their union. A ratio of 0 means a perfect overlap of two regions, and 1 means completely dissimilarity. This ratio for m RoIs is defined as [41]:

$$\vartheta(S) = \prod_{i=1}^{m} \left(\frac{|R_i \cup R_i(S)| - |R_i \cap R_i(S)|}{|R_i \cup R_i(S)|} \right)^{\frac{1}{m}} \tag{9}$$

where S is the 3D bone model, R_i is the set of pixels in i^{th} RoI in segmented image, $R_i(S)$ is the set of pixels of the projection of i^{th} RoI in 3D model S, and $|R|$ indicates the number of pixels in region R (see Fig. 10).

The likelihood energy term is defined as

$$E_l(S) = (1 - \xi(S)) + \vartheta(S) \tag{10}$$

Table 1 Correlations of patella, talus, and pelvis at different reduced dimensions. The higher the reduced dimension, the lesser information is lost in the data

	$\mathcal{M} = 1$	$\mathcal{M} = 2$	$\mathcal{M} = 3$
Patella	0.753	0.84	0.896
Talus	0.847	0.92	0.94
Pelvis	0.682	0.734	0.819

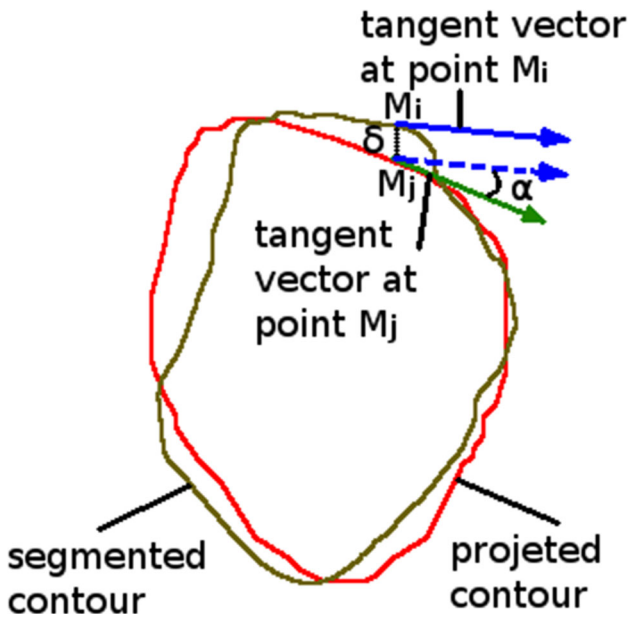


Fig. 9 Directional component used in Eq. (7)

2.4.2 Prior energy term

In order to favor statistically admissible surface structures in our training base, our prior energy term will favor a candidate shape in a high-density area of our shape density mapping and will be simply defined as the average distance between it and its k nearest neighbors:

$$E_p(s) = \frac{1}{k} \sum_{i}^{k} (s - s_i)^2 \tag{11}$$

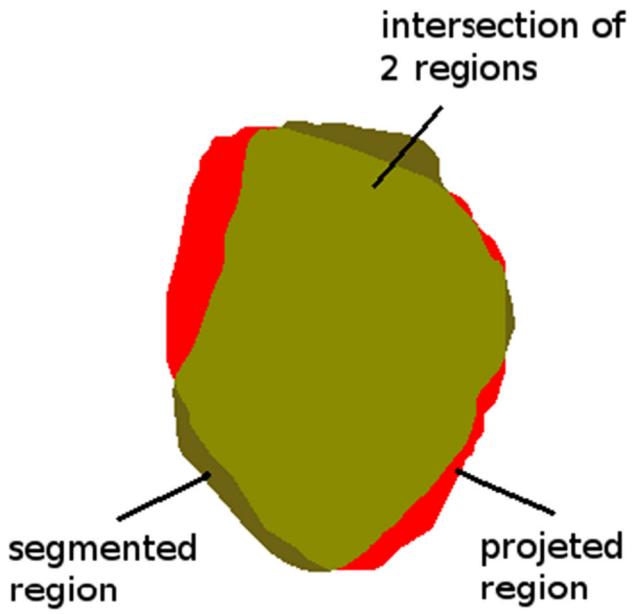


Fig. 10 Overlap of segmented and projected regions

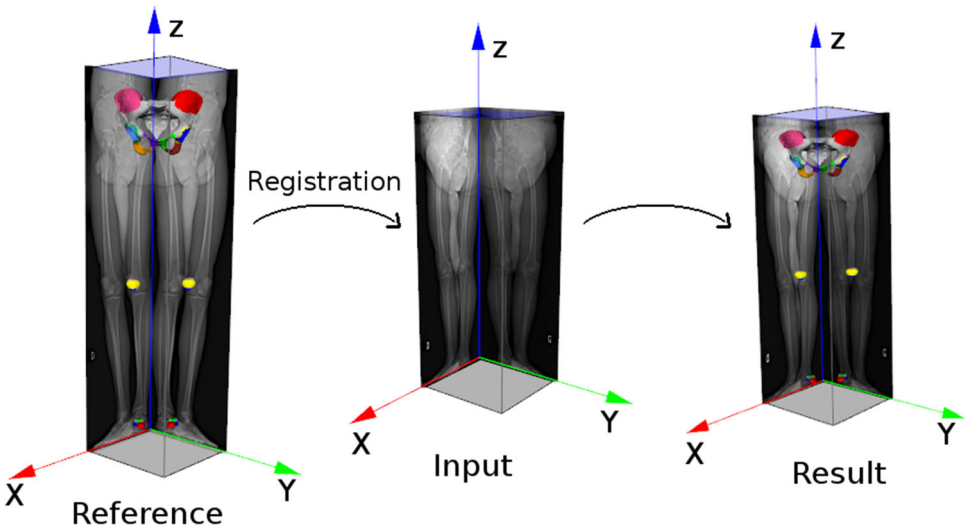


Fig. 11 Initialization step is the registration of the reference box on the input box

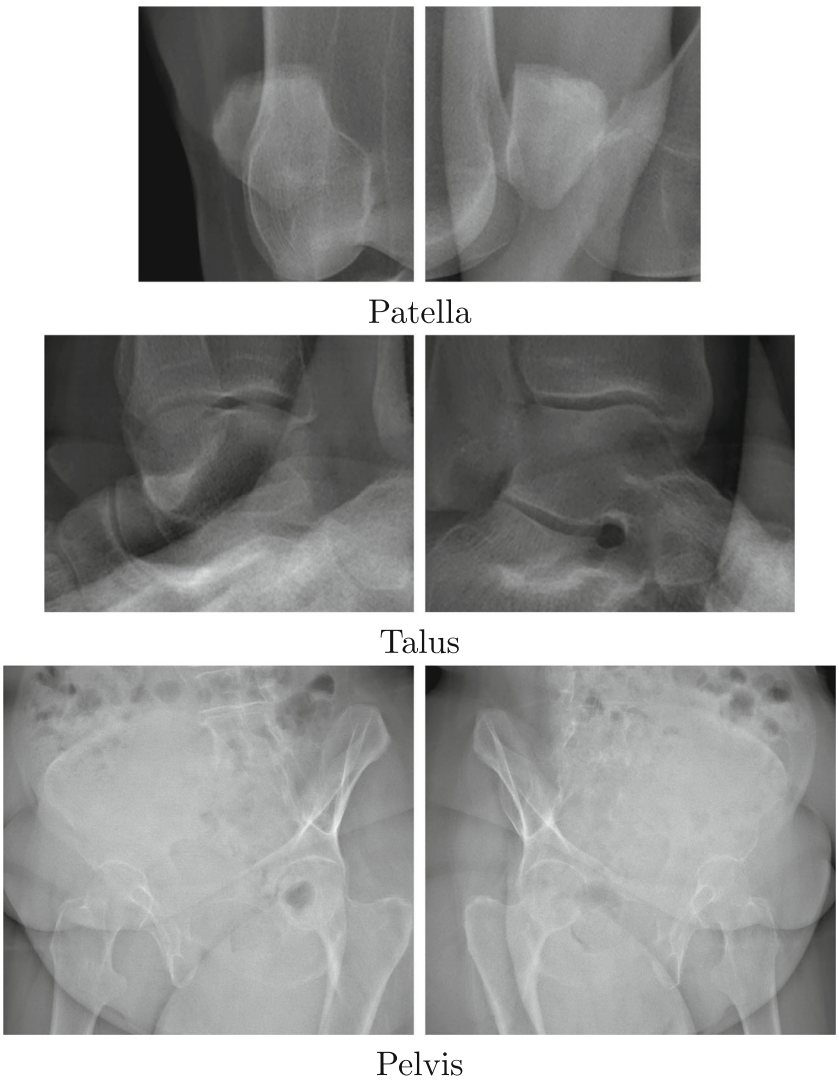


Fig. 12 Example of bone images after initialization (see Sect. 2.5.1) step. The 45-degree view is on the left, and 135-degree view is on the right of each bone

Table 2 Results on comparison of 5 patellae

Model	RMS (mm)	Maximum error (mm)	Errors < 1 mm (%)	1 mm ≤ Errors < 2 mm (%)	Errors ≥ 2 mm (%)
1	0.8	2.6	81.4	18.2	0.4
2	0.9	3.3	74.9	21.9	3.2
3	0.6	2.0	89.7	10.3	0.0
4	1.0	3.0	69.6	26.7	3.7
5	1.3	4.1	59.9	28.4	11.7
Average ± SD	0.9 ± 0.2	3.0 ± 0.7	75.1 ± 10.1	21.1 ± 6.5	3.8 ± 4.2

where s_i is a point in the k nearest neighbourhood of s in the reduced dimensionality space.

2.5 3D reconstruction

2.5.1 Initialization

From biplanar oblique (45 and 135 degree) X-ray images (I_{45} and I_{135}), we construct a box which contains the bone to be reconstructed. The dimensions of the box are the width of the 45-degree image, the width of the 135-degree image, and the height of the highest image of the bone. Then, we manually position a 3D model of the bone (patella, talus, pelvis) in this box in such a way that the projections of these 3D models coincide with their corresponding bone in the two images. This ref_{box} box constitutes a reference box, and these bone models ($ref_{patella}, ref_{talus}, ref_{pelvis}$) constitute reference models. These references are computed offline and only once.

The initialization step of our approach is simply the registration of ref_{box} on in_{box} which is created by using the aforementioned method on biplanar input images. The transformation from this registration is then applied to the ref_{box} and the reference models ($ref_{patella}, ref_{talus}, ref_{pelvis}$). Figure 11 well illustrates the initialization step, and Fig. 12 shows a result example. This step can be refined by using a 3D/2D registration. The registered reference models will be used for the next step.

2.5.2 Optimization

Our method is based on the optimization of an objective function E that contains both a contour/region-based similarity term E_l and a shape prior term E_p:

$$E(S(\theta)) = E_l(S, I_{45}) + E_l(S, I_{135}) + \beta E_p(s) \tag{12}$$

where θ is the deformation parameters, S and s are the bone shape in the original and reduced dimensionality space, and β is a weighting factor which indicates the importance of the prior information.

To minimize this complex non-convex energy function, we resort to Exploration Selection (ES) algorithm [42], a stochastic and efficient optimization algorithm. This algorithm belongs to the class of evolutionary algorithms and is typically well-suited for this type of function to be optimized. This class of algorithm is inspired by the evolution of nature by reflecting the process of natural selection to provide good approximate solutions to complex problems. These algorithms have been successfully used in medical imaging [14, 43, 44].

The ES algorithm can be summarized in two steps: exploration and selection steps. Let F be a finite discrete subset of the Cartesian product of h intervals $[m_i, M_i]$ for $1 \leq i \leq h$, and $\theta = \{\theta_1, ..., \theta_n\}$ a set of n candidate solutions (population) randomly chosen. $\hat{\theta}$ is the optimal element θ_i of θ such that $E(\theta_j) > E(\theta_i)$ for $1 \leq j < i$, and $E(\theta_j) \geq E(\theta_i)$ for $i < j \leq n$. In the first step, each solution of θ is considered as an individual that attempts a random search on the exploration graph. The exploration process acts independently on each individual and chooses a random f according to a positive distribution. We compute $\hat{\theta} = \arg\min_\theta E(.)$, and for $i \leq f$, we replace θ_i by $\gamma_i \in N(\theta_i) \setminus \{\hat{\theta}\}$ according to a uniform distribution, where $N(b)$ is defined by $\{a \in F$: for some $j, \mid a_j - b_j \mid \leq r(M_j - m_j), a_i = b_i, i \neq j\}$, and $r \in [0, 1]$ is the radius of exploration. Otherwise, θ_i is changed by $\hat{\theta}$ in the second step. This process is run until a stopping criterion has been met (see Algorithm 3).

Table 3 Results on comparison of 5 tali

Model	RMS (mm)	Maximum error (mm)	Errors < 1 mm (%)	1 mm ≤ Errors < 2 mm (%)	Errors ≥ 2 mm (%)
1	1.7	6.5	48.1	29.8	22.1
2	1.2	3.9	56.4	36.4	7.2
3	1.3	4.4	57.8	31.2	11.0
4	1.5	4.8	44.1	39.8	16.1
5	1.5	3.9	51.5	31.3	17.2
Average ± SD	1.4 ± 0.2	4.7 ± 1.0	51.6 ± 5.1	33.7 ± 3.8	14.7 ± 5.2

Table 4 Results on comparison of 3 pelvis

Model	RMS (mm)	Maximum error (mm)	Errors < 1mm (%)	1mm ≤ Errors < 2mm (%)	2mm ≤ Errors < 4mm (%)	Errors ≥ 4mm
1	3.6	14.3	25.8	24.9	25.0	24.3
2	2.5	13.9	40.0	25.6	22.8	11.6
3	4.2	19.3	20.8	21.0	29.3	28.9
Average ± SD	3.4 ± 0.7	15.8 ± 2.5	28.9 ± 8.1	23.8 ± 2.0	25.7 ± 2.7	21.6 ± 7.3

Algorithm 3 ES optimization algorithm

$E(.)$	A h-variable function defined on F to be minimized
$N(a)$	The neighborhood of an element $a \in F$ defined by $\{b \in F$: for some $1 \leq j \leq l$, $\|b_j - a_j\| \leq r\ (M_j - m_j)$, $b_i = a_i,\ i \neq j\}$
D	$(= h/r)$ The diameter of the exploration graph F (endowed with the neighborhood system $\{N(a)\}_{a \in F}$)
θ	$\theta = (\theta_1, \ldots, \theta_n)$, an element of F^n (population)
$\hat{\theta}$	$\hat{\theta} \in F$, $\hat{\theta} = \arg\min_{\theta_i \in (\theta_1, \ldots, \theta_n)} E(\theta_i)$ *i.e.*, the minimal point in θ with the lowest label
p	The probability of exploration
t	The iteration step
F	A finite discrete subset of the Cartesian product $\Pi_{j=1}^{h}[m_j, M_j]$ of h compact intervals
n	The size of the population (greater than D)
r	A real number $\in [0, 1]$ called the radius of exploration (with r greater than the ϵ-machine)

1. Initialization
Random initialization of $\theta = (\theta_1, \ldots, \theta_n) \in F^n$
$t \leftarrow 2$

2. Exploration/Selection
while a stopping criterion is not met **do**

1. Compute $\hat{\theta}$; $\hat{\theta} \leftarrow \arg\min_{\theta_i \in \theta} E(\theta_i)$
2. Draw f according to the binomial law $b(n, p)$
   - For $i \leq f$, replace θ_i by $\gamma_i \in N(\theta_i) \setminus \{\hat{\theta}\}$ according
   - to the uniform distribution (*Exploration* step)
   - For $i > f$, replace θ_i by $\hat{\theta}$ (*Selection* step)
3. $t \leftarrow t + 1$ and $p \leftarrow t^{-1/D}$

end while

2.5.3 Refinement

We use an additional strategy which consists in associating the stochastic exploration, selection search with a local optimization technique. In each generation (i.e., each Exploration/Selection step), the best individual or solution is used to initialize a gradient descent-style deterministic minimization technique in which local deformations are used to refine the bone shape result previously estimated by the stochastic optimizer (at each E/S step). Therefore, the best individual deterministically explores the local neighborhood in parameter space to find a point that further minimizes our energy function. More precisely, this deterministic gradient descent procedure uses local deformations to refine the bone shape obtained with θ previously estimated. This procedure can be described as follows:

- Do
  - For each point p_i on external contour of bone shape $S(\theta)$
    - Find K_i nearest neighbour of p_i
    - Compute the normal η_i which is the average of the normal of K_i
    - Compute $\vartheta(S(\theta, K_i))$, $\vartheta(S(\theta, K_i - \varepsilon\eta))$, $\vartheta(S(\theta, K_i + \varepsilon\eta))$
    - Update K_i with the lowest energy ϑ
- While the energy ϑ is not stable.

3 Experiments

3.1 Lower limb (patella, talus, pelvis) database

Our training databases consist of 654 surface models for the patella, 380 for the talus, and 39 for the pelvis. These training surface models were constructed from binary volumes that were semi-automatically segmented from CT-datasets of patient [27].

3.2 Radiographic images

In our application, we used biplanar oblique (45 and 135 degree) X-ray images (I_{45} and I_{135}) acquired with a low dose imaging device. This system performs the simultaneous acquisition of two orthogonal calibrated X-rays, with the patient in a standing position. The radiographic image illustrates the superposition of the various structures on the same plan and gives place to semi-transparent images. The size of our radiographic images is approximately 1764×5932 pixels (coded on 256 gray levels).

3.3 Comparison protocol

We only had 13 bones (5 patellae, 5 tali, and 3 pelvis) from 13 patients (13 pairs of radiographic images (I_{45} and I_{135} views) of the lower limbs) to validate our 3D reconstruction method. This comparison was made using the distance (root mean square (RMS) and maximum) between the point from the reconstructed lower limb and the nearest point of the corresponding lower limb obtained with CT-scan, which was considered as the ground truth and whose accuracy is ±1 mm for the human spine [45].

3.4 Experimental results

In our study, we use the shape density mapping, obtained by the FastMap, in the reduced space of dimension 3, that preserves 94%

information for talus and more than 80% information for pelvis and patella in terms of distance of object pairs.

Based on preliminary test results, we choose k = 12, 8, 18 for the nearest neighbour and ϱ = 1.2, 2.8, 0.8 for the power parameter in the IDW interpolation for the patella, talus, and pelvis, respectively. The prior weighting factor β was set to 0.001, 0.0005, and 0.00005 for the patella, talus, and pelvis, respectively, to weight the prior energy term with respect to the likelihood energy term. Finally, we set the size of the population to 100 and the number of iterations to 380 for ES algorithm.

We have validated our method on 5 patellae, 5 tali, and 3 pelvis by using the aforementioned protocol. The results of comparisons are given in Tables 2, 3, and 4. Figure 13 details the errors of each model of each bone. Figures 14, 15, and 16 show examples of the reconstructed bone by our method and their projections on both images.

For the patella, the average RMS was 0.9 mm and the average maximum error was 3.0 mm. 75.1% of points had an error less than 1 mm. 21.1% of points had an error between 1 and 2 mm. The error of the remaining 3.8% of points was greater than 2 mm.

For the talus, the average RMS was 1.4 mm and the average maximum error was 4.7 mm. 51.6% of points had an error less than 1 mm. 33.7% of points had an error between 1 and 2 mm. The error of the remaining 14.7% of points was greater than 2 mm.

For the pelvis, the average RMS was 3.4 mm and the average maximum error was 15.8 mm. 52.7% of points had an error less than 2 mm. 25.7% of points had an error between 2 and 4 mm. The error of the remaining 21.6% of points was greater than 4 mm.

4 Discussion

In this paper, we present a new approach for the 3D reconstruction of human lower limbs (patella, talus, and pelvis) from two radiographic oblique projections (45 and 135 degrees). This approach efficiently exploits all the information existing in the image (edges and RoIs) and the estimation of a FastMap both for the reduction of the dimensionality and to constrain the 3D reconstruction problem. This 3D reconstruction problem is considered as a cost function optimization problem encompassing the information extracted from the image and the information extracted from the database by FastMap. Our 3D reconstruction method has the advantage of minimizing the X-ray dose (only two oblique radiographic images), exploitation of the maximum amount of information existing in these two images (contours and RoIs), and the use of the nonlinear FastMap statistical method which gives a good representation of deformations in the bone database relative to those used in other 3D reconstruction methods. The results obtained in our experiments are encouraging and indicate that accurate unsupervised 3D reconstruction is technically feasible.

This approach has been validated on a sample of 13 bones (5 patellae, 5 tali, and 3 pelvis) from 13 patients (13 pairs of radiographic images (I_{45} and I_{135} views) of lower limbs), by comparing the model obtained from our approach and those obtained with CT-scan, which was considered as the ground truth and whose accuracy is $\pm$1 mm for the human spine [45]. The average RMS was 0.9 mm, 1.4 mm, and 3.4 mm for the patella, the talus, and the pelvis, respectively. The mean and the standard deviation of the percentages for the points whose error is less than 1 mm for the patella and the talus were $75.1 \pm 10.1\%$ and $51.6 \pm 5.1\%$. Those whose error is less than 2 mm for the pelvis were $52.7 \pm 9.8\%$.

Our results are better than that of the hierarchical statistical modeling method [14] in case of the patella and talus (see Table 5). The errors

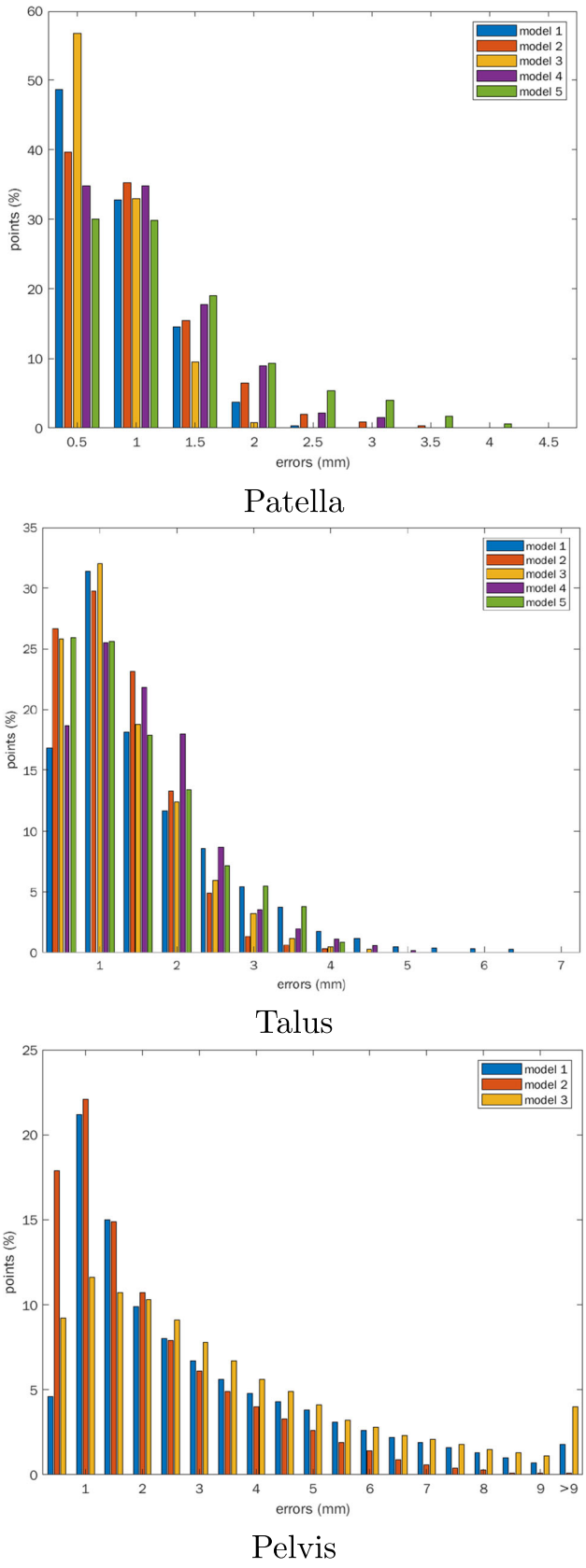


Fig. 13 Detailed errors of each model of each bone

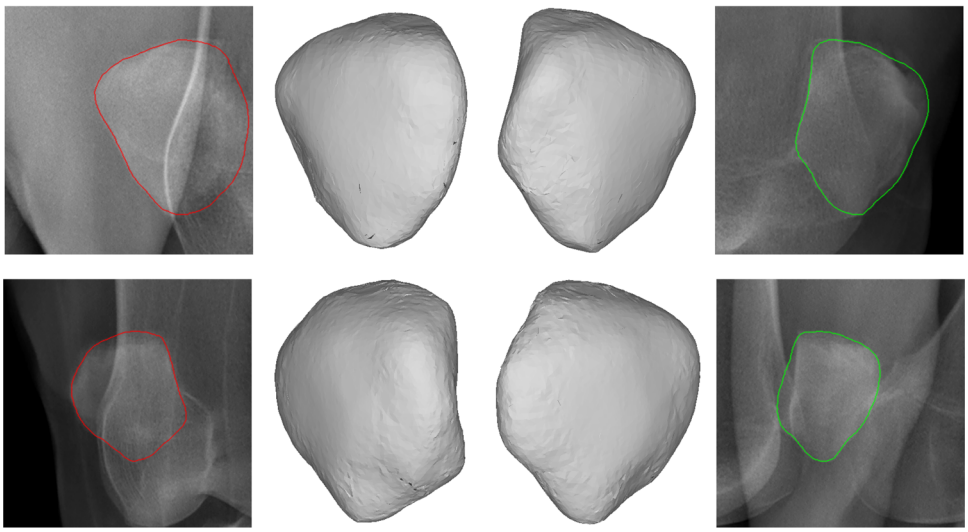


Fig. 14 Examples of the reconstructed patellae by our method and their projections on both images

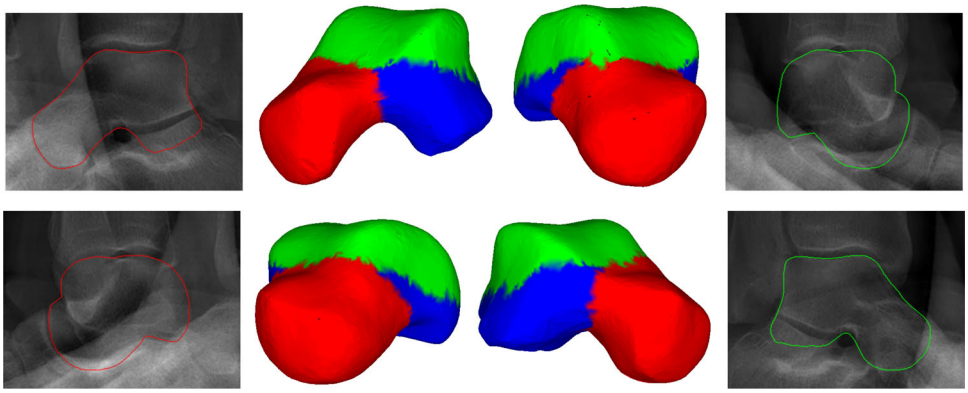


Fig. 15 Examples of the reconstructed tali by our method and their projections on both images

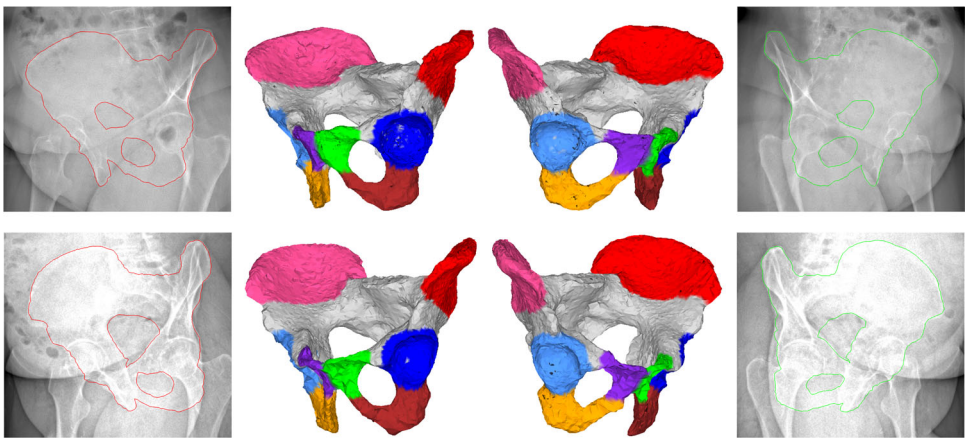


Fig. 16 Examples of the reconstructed pelvis by our method and their projections on both images

Table 5 Results of our method versus the hierarchical statistical modeling method [14]

Our method			**Hierarchical statistical modeling method**		
Bones	**Average RMS** (mm)	**Average maximum error** (mm)	**Bones**	**Average RMS** (mm)	**Average maximum error** (mm)
Patella	0.9	3.0	**Thoracic vertebra**	1.6	4.5
Talus	1.4	4.7	**Lumbar vertebra**	1.9	5.4
Pelvis	3.4	15.8			

of the pelvis reconstruction were greater than other bone reconstructions because the pelvis has a very complex structure and its database was smaller than that of the other bones, limited to 39 objects in our application. Also, the sacrum and the coccyx were blurred in the biplanar images, and the L5 S1 junction is hard to identify. Note that the structural complexity, the anatomical position in the human body, and the number of elements in the training set of each bone of each method are different.

Compared to the biplanar method [14] which uses only the contours of anatomical shapes, our likelihood function uses all the information of the two images (contours and RoIs). Moreover, the linear PCA-based dimensionality reduction used in [14] does not perfectly represent admissible statistical deformations existing in the training set because linearity is a hypothesis that is not necessarily true in our context. However, our statistical deformable model requires the selection of the number of the nearest neighbor (k) to generate a new instance. This can be a limitation as it can be difficult to determine the optimal value of k for a given dataset.

The biplanar technique [14] is also limited due to the inherent inaccuracy produced in the segmentation of I_{LAT} (leading to reconstruction errors). In addition, this method does not use all the information contained in two X-ray projections, for example, the contours of each bone structure and the geometric structure or statistical knowledge of the possible deformation of the bone structure to be reconstructed.

Our approach requires a training representative database. Nevertheless, our proposed reconstruction method remains unsupervised in the sense that this database is constructed *off-line* and not during the 3D reconstruction step.

5 Conclusion

Our approach provides an accurate representation of the patella, talus, and pelvis from just two X-ray views, while the CT-scan requires hundreds of images to achieve the same three-dimensional reconstruction. Our proposed method is interesting in terms of the quantity of data to be acquired, processed, and managed.

Our technique constitutes an alternative to CT-scan 3D reconstruction with the advantage of low radiation. This promises to be of great interest for the study of bone deformities, simulation of orthopedic treatments, and for reliable geometric models for finite element studies. Presently, this reconstruction method is not suitable without improvement for surgical navigation applications as compared to the current standard CT-scan reconstruction with errors of ± 1 mm.

To our knowledge, our proposed approach is the first non-CT 3D reconstruction method for the talus. Our technique is sufficiently robust to be applied to other medical reconstruction problems for which a database of the anatomical structure is available. In addition, our method, based on the ES optimizer, though slightly time-consuming, is easily parallelizable and thus remains especially well-suited for the next-generation GPU or massively parallel computers and multi-core processors. Furthermore, FastMap algorithm uses a distance metric to transform data from a high-dimensional space into a low-dimensional space. This means that another distance metrics (other than Euclidean distance) can be applied.

Acknowledgements We would like to thank Nicholas Newman MD FRCSC for english language editing.

References

1. Terzopoulos D, Witkin A, Kass M (1988) Constraints on deformable models: recovering 3D shape and nonrigid motion. Artificial Intelligence 36(1):91–123
2. Benjamin R, Prakoonwit S, Matalas I, Kitney R (1996) Object-based three-dimensional X-ray imaging. Medical and Biological Engineering and Computing 34(6):423–430
3. Nikkhade-Dehkordi B, Bro-Nielsen M, Darvann T, Gramkow C, Egund N, Hermann K (1996) 3D reconstruction of the femoral bone using two X-ray images from orthogonal views. In: Computer Assisted Radiology, p 1015
4. Kita Y (1996) Elastic-model driven analysis of several views of a deformable cylindrical object. IEEE Trans Pattern Anal Mach Intell 18(12):1150–1162
5. Bardinet E, Cohen LD, Ayache N (1998) A parametric deformable model to fit unstructured 3D data. Comp Vision Image Underst. 71(1):39–54
6. Mitton D, Landry C, Veron S, Skalli W, Lavaste F, De Guise JA (2000) 3D reconstruction method from biplanar radiography using non-stereo corresponding points and elastic deformable meshes. Med Biol Eng Comput 38(2):133–139
7. Baka N, Kaptein BL, de Bruijne M, van Walsum T, Giphart J, Niessen WJ, Lelieveldt BP (2011) 2D–3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. Med Image Anal 15(6):840–850
8. Ehlke M, Ramm H, Lamecker H, Hege HC, Zachow S (2013) Fast generation of virtual X-ray images for reconstruction of 3D anatomy. IEEE Trans Vis Comput Graph 19(12):2673–2682
9. Zhang J, Besier TF (2017) Accuracy of femur reconstruction from sparse geometric data using a statistical shape model. Computer methods in biomechanics and biomedical engineering 20(5):566–576
10. Fotsin TJT, Vázquez C, Cresson T, De Guise J (2019) Shape, pose and density statistical model for 3D reconstruction of articulated structures from X-ray images. In: 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2748–2751
11. Gajny L, Girinon F, Bayoud W, Lahkar B, Bonnet-Lebrun A, Rouch P, Lazennec JY, Skalli W (2022) Fast quasi-automated 3D reconstruction of lower limbs from low dose biplanar radiographs using statistical shape models and contour matching. Medical Engineering & Physics 101
12. Reyneke CJF, Lüthi M, Burdin V, Douglas TS, Vetter T, Mutsvangwa TE (2018) Review of 2-D/3-D reconstruction using statistical shape and intensity models and X-ray image synthesis: toward a unified framework. IEEE Rev Biomed Eng 12:269–286
13. Kim H, Lee K, Lee D, Baek N (2019) 3D reconstruction of leg bones from X-ray images using CNN-based feature analysis. In: International Conference on Information and Communication Technology Convergence (ICTC), pp 669–672
14. Benameur S, Mignotte M, Labelle H, De Guise JA (2005) A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. IEEE Trans Biomed Eng 52(12):2041–2057
15. Benameur S, Mignotte M, Destrempes F, De Guise JA (2005) Three-dimensional biplanar reconstruction of scoliotic rib cage

using the estimation of a mixture of probabilistic prior models. IEEE Trans Biomed Eng 52(10):1713–1728
16. Yang CJ, Lin CL, Wang CK, Wang JY, Chen CC, Su FC, Lee YJ, Lui CC, Yeh LR, Fang YHD (2022) Generative adversarial network (GAN) for automatic reconstruction of the 3D spine structure by using simulated bi-planar X-ray images. Diagnostics 12(5):1121
17. Kasten Y, Doktofsky D, Kovler I (2020) End-to-end convolutional neural network for 3D reconstruction of knee bones from bi-planar X-ray images. In: International Workshop on Machine Learning for Medical Image Reconstruction, pp 123–133
18. Ahishakiye E, Van Gijzen MB, Tumwiine J, Wario R, Obungoloch J (2021) A survey on deep learning in medical image reconstruction. Intelligent Medicine
19. Maken P, Gupta A (2022) 2D-to-3D: a review for computational 3D image reconstruction from X-ray images. Archives of Computational Methods in Engineering, p 1–30
20. Binol H, Niazi MKK, Plotner A, Sopkovich J, Kaffenberger BH, Gurcan MN (2020) A multidimensional scaling and sample clustering to obtain a representative subset of training data for transfer learning-based rosacea lesion identification. In: Medical Imaging: Computer-Aided Diagnosis, International Society for Optics and Photonics 11314:1131415
21. Cao X, Chen C, Tian L (2021) Supervised multidimensional scaling and its application in MRI-based individual age predictions. Neuroinformatics 19(2):219–231
22. Nguyen DCT, Benameur S, Mignotte M, Lavoie F (2018) Superpixel and multi-atlas based fusion entropic model for the segmentation of X-ray images. Med Image Anal 48:58–74
23. Zhang YJ (2006) An overview of image and video segmentation in the last 40 years. Advances in Image and Video Segmentation, p 1–16
24. Cufi X, Munoz X, Freixenet J, Marti J (2006) A review of image segmentation techniques integrating region and boundary information. Advances in imaging and electron physics 120:1–39
25. Muñoz X, Freixenet J, Cufı X, Martı J (2003) Strategies for image segmentation combining region and boundary information. Pattern recognition letters. 24(1–3):375–392
26. Faloutsos C, Lin KI (1995) FastMap: a fast algorithm for indexing, data-mining and visualization of traditional and multimedia datasets. In: Proceedings of the 1995 ACM SIGMOD International Conference on Management of Data, pp 163–174
27. Huang TC, Zhang G, Guerrero T, Starkschall G, Lin K-P, Forster K (2006) Semi-automated CT segmentation using optic flow and Fourier interpolation techniques. Computer methods and programs in biomedicine 84(2–3):124–134
28. Rusinkiewicz S, Levoy M (2001) Efficient variants of the ICP algorithm. In: Proceedings Third International Conference on 3-D Digital Imaging and Modeling, pp 145–152
29. Latulippe M (2013) Calage robuste et accéléré de nuages de points en environnements naturels via l'apprentissage automatique
30. Buades A, Coll B, Morel JM (2011) Non-local means denoising. Image Processing On Line 1:208–212
31. Canny J (1986) A computational approach to edge detection. IEEE Trans pattern Anal Mach Intell. 6:679–698
32. Chen H, Bhanu B (2008) Efficient recognition of highly similar 3D objects in range images. IEEE Trans Pattern Anal Mach Intell. 31(1):172–179
33. Touati R, Mignotte M (2014) MDS-based multi-axial dimensionality reduction model for human action recognition. In: 2014 Canadian Conference on Computer and Robot Vision, pp 262–267
34. Belhadj LC, Mignotte M (2016) Spatio-temporal fastmap-based mapping for human action recognition. In: 2016 IEEE International Conference on Image Processing (ICIP), pp 3046–3050
35. Mignotte M (2011) MDS-based multiresolution nonlinear dimensionality reduction model for color image segmentation. IEEE transactions on neural networks 22(3):447–460
36. Mignotte M (2012) MDS-based segmentation model for the fusion of contour and texture cues in natural images. Comput Vis Image Underst. 116(9):981–990
37. Touati R, Mignotte M (2016) A multidimensional scaling optimization and fusion approach for the unsupervised change detection problem in remote sensing images. In: Sixth International Conference on Image Processing Theory, Tools and Applications (IPTA), pp 1–6
38. Jacobson NP, Gupta MR (2005) Design goals and solutions for display of hyperspectral images. IEEE Transactions on Geoscience and Remote Sensing 43(11):2684–2692
39. Cox MA, Cox TF (2008) Multidimensional scaling. In: Handbook of Data Visualization, p 315–347
40. Jain AK, Zhong Y, Lakshmanan S (1996) Object matching using deformable templates. IEEE Trans pattern Anal Mach I ntell. 18(3):267–278
41. Jáuregui DAG, Horain P (2009) Region-based vs. edge-based registration for 3D motion capture by real time monoscopic vision. In: International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications, pp 344–355
42. François O (2002) Global optimization with exploration/selection algorithms and simulated annealing. The Annals of Applied Probability 12(1):248–271
43. Pena-Reyes CA, Sipper M (2000) Evolutionary computation in medicine: an overview. Artif Intell Med. 19(1):1–23
44. Mignotte M, Meunier J, Tardif JC (2001) Endocardial boundary e timation and tracking in echocardiographic images using deformable template and markov random fields. Pattern Analysis & Applications 4(4):256–271
45. Aubin CE, Dansereau J, Parent F, Labelle H, de Guise JA (1997) Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine. Medical and biological engineering and computing 35(6):611–618

Dac Cong Tai Nguyen received an M.Sc. degree from Claude Bernard University Lyon 1, France, and a Ph.D. degree in computer science from the University of Montreal (DIRO), Canada. His research interests concern image processing, computer vision, 3D reconstruction, and its applications to medical imaging.

Said Benameur obtained a computer engineering degree from Mouloud Mammeri University, Algeria (1985-1990); an M.Sc. degree in computer science from the University of Quebec in Montreal (UQAM), Canada (1996-1998), and a Ph.D. in engineering from the School of Engineers (ETS), University of Quebec, Canada in (2000-2004). He was a post-doctoral fellow at the University of Montreal (DIRO), Canada (Quebec), from April 2006 to April 2008. His research interests include statistical learning, image processing and computer vision.

Max Mignotte received a D.E.A. degree in digital signal, image, and speech processing from the Grenoble Institute of Technology, France, and a Ph.D. degree in electronics and computer engineering from the University of Western Brittany, France, and the Digital Signal Laboratory, French Naval Academy, France. His research interests include statistical methods, Bayesian inference, and hierarchical models for high-dimensional inverse problems, such as segmentation, parameter estimation, fusion, shape recognition, deconvolution, 3D reconstruction, and restoration problems.

Frédéric Lavoie received an M.D. degree from Laval University in Quebec City and an M.Sc. degree from the University of Montreal, Canada, and completed his residency in orthopaedics at the University of Montreal, Canada. His research interests include 3D imaging of the lower limbs, 3D kinematics of the knee, and computer-assisted surgery.

Authors and Affiliations

Dac Cong Tai Nguyen[1,2] · Said Benameur[2] · Max Mignotte[1] · Frédéric Lavoie[2,3]

Said Benameur
benameur.said@gmail.com

Max Mignotte
mignotte@iro.umontreal.ca

Frédéric Lavoie
flavoie@eiffelmedtech.com

[1] Département d'Informatique et de Recherche Opérationnelle (DIRO), Université de Montréal, Québec Montréal, Canada

[2] Eiffel Medtech Inc.,, Québec Montréal, Canada

[3] Orthopedic Surgery Department, Centre Hospitalier de l'Université de Montréal (CHUM),, Québec Montréal, Canada