Medical Image Analysis 52 (2019) 42–55

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Supervised learning for bone shape and cortical thickness estimation from CT images for finite element analysis

Vimal Chandran[a,*], Ghislain Maquer[a], Thomas Gerig[b], Philippe Zysset[a], Mauricio Reyes[a,*]

[a] Institute of Surgical Technology and Biomechanics, University of Bern, Switzerland
[b] Graphics and Vision Research Group, University of Basel, Switzerland

ARTICLE INFO

Article history:
Received 19 January 2018
Revised 24 August 2018
Accepted 2 November 2018
Available online 16 November 2018

Keywords:
Hip fracture
Cortical thickness
Shape regression
Super resolution
Finite elements
Gaussian process model

ABSTRACT

Knowledge about the thickness of the cortical bone is of high interest for fracture risk assessment. Most finite element model solutions overlook this information because of the coarse resolution of the CT images. To circumvent this limitation, a three-steps approach is proposed. 1) Two initial surface meshes approximating the outer and inner cortical surfaces are generated via a shape regression based on morphometric features and statistical shape model parameters. 2) The meshes are then corrected locally using a supervised learning model build from image features extracted from pairs of QCT (0.3-1 mm resolution) and HRpQCT images (82 µm resolution). As the resulting meshes better follow the cortical surfaces, the cortical thickness can be estimated at sub-voxel precision. 3) The meshes are finally regularized by a Gaussian process model featuring a two-kernel model, which seamlessly enables smoothness and shape-awareness priors during regularization. The resulting meshes yield high-quality mesh element properties, suitable for construction of tetrahedral meshes and finite element simulations. This pipeline was applied to 36 pairs of proximal femurs (17 males, 19 females, 76 ± 12 years) scanned under QCT and HRpQCT modalities. On a set of leave-one-out experiments, we quantified accuracy (root mean square error $= 0.36 \pm 0.29\,mm$) and robustness (Hausdorff distance $= 3.90 \pm 1.57\,mm$) of the outer surface meshes. The error in the estimated cortical thickness ($0.05 \pm 0.40\,mm$), and the tetrahedral mesh quality (aspect ratio $= 1.4 \pm 0.02$) are also reported. The proposed pipeline produces finite element meshes with patient-specific bone shape and sub-voxel cortical thickness directly from CT scans. It also ensures that the nodes and elements numbering remains consistent and independent of the morphology, which is a distinct advantage in population studies.

1. Introduction

Hip fracture is a common injury among the elderly population (Cauley, 2013). Associated with high morbidity and mortality rates, such incident in hip fractures are particularly debilitating, often leads to a loss of independence and are responsible for substantial health-care costs (Cauley et al., 2016). Aside from alterations of the trabecular structure and density (Taghizadeh et al., 2017), mounting evidence suggests that the fracture risk is also largely amplified by the local thinning of the cortex resulting from osteoporosis (Mayhew et al., 2005; Poole et al., 2017). Unfortunately, the current diagnosis of osteoporosis rests on DXA (dual X-rays absorptiometry), which is bi-dimensional, includes indistinctly cortical and trabecular bone regions and sometimes results in patient being erroneously considered not at risk (Stone et al., 2003). Alternatively, quantitative computer tomography (QCT) provides distinct 3D information for each bone region. Performed at the hip, QCT is an alternative diagnostic tool (Engelke et al., 2015) that can also be used for generating patient-specific finite element (FE) models (Zysset et al., 2015). FE models are computer simulations used to perform mechanical tests virtually. They allow a non-invasive evaluation of bone strength, which is the force required to fracture the bone. The Food and Drug Administration cleared the use of FEA-derived bone strength calculations (VirtuOst, O.N.Diagnostics, Berkeley, CA, USA) as a class II medical device for estimating fracture risk assessment and monitoring treatment efficacy (Keaveny, 2017). However, several technical challenges hinder direct application of FEA for clinical use. Main factors are summarized below, and represent the main motivation of the presented work.

Generating a patient-specific FE model with accurate geometry and cortical thickness is a complex engineering task. The standard work-flow starts with the segmentation of both trabecular and cortical regions. Ex vivo studies showed that this step can be done

* Corresponding authors.
E-mail addresses: vimalc@ieee.org (V. Chandran), mauricio.reyes@istb.unibe.ch (M. Reyes).

https://doi.org/10.1016/j.media.2018.11.001
1361-8415/

accurately from high-resolution peripheral QCT (HRpQCT, 61-82 µm voxel size) (Pahr and Zysset, 2009). However on QCT images, the segmentation is prone to errors because of partial volume effects arising from the coarser voxel size. In that case, there is no clear boundary between bone regions (Peleg et al., 2014). The cortex is often thinner than the point spread function of the CT scanner and may not even be visible (Treece et al., 2012). The finite element mesh is then generated from the segmentation masks, which are also challenging because of the thin cortex, and the need to yield a volumetric mesh guaranteeing an appropriate number of elements and element quality for FE analysis to converge (Arbenz and Flaig, 2008; Liu et al., 2007).

Despite previous contributions to bone strength (Luisier et al., 2014) and fracture risk analysis (Treece et al., 2015), most FE models do not include a realistic cortical shell (Bucki et al., 2010; Pauchard et al., 2016; Maquer et al., 2016). This is even more surprising considering the amount of studies that have focused on measuring cortical thickness from QCT images (Prevrhal et al., 1999; 2003; Pakdel et al., 2012; Treece et al., 2010; 2012; Humbert et al., 2016; Treece and Gee, 2015). Except from (Carballido-Gamio et al., 2015), the lack of a unified framework combining accurate FE modelling and cortical thickness estimation hinders the adoption of FE-based diagnosis into clinical routine of osteoporosis (Engelke et al., 2016; Zysset et al., 2015).

Accordingly, the aim of this work is to develop a framework for generating high quality meshes for bones from QCT images including a cortical thickness estimated at sub-voxel accuracy. To this end we present a three-step approach: 1) Two initial surface meshes approximating the outer and inner cortical surfaces are generated via a shape regression based on morphometric features and statistical shape model parameters. 2) The meshes are then corrected locally using a supervised learning model build from image features extracted from pairs of QCT (0.3-1 mm resolution) and HRpQCT images (82 µm resolution). As the resulting meshes better follow the cortical surfaces, the cortical thickness can be estimated at sub-voxel precision. 3) The meshes are finally regularized by a Gaussian process model featuring a two-kernel model, which seamlessly enables smoothness and shape-awareness priors during regularization. We demonstrate its performance on a leave-one-out (LOO) study based on an ex vivo database of 72 proximal femurs scanned with QCT and HRpQCT modalities.

The paper is organised as follows, in Section 2 we outline the proposed framework. In Section 3, dataset, experimental set-up and results are presented. In Section 4 results are discussed and finally, Section 5 ends conclusions and outlook.

2. Methodology

An overview of the proposed approach is given in Fig. 1. Each step is described in details in the following sections.

2.1. Step 1: Coarse shape regression

Shape regression using supervised learning have already been proposed (Zhou and Comaniciu, 2007; Lindner et al., 2015). In those studies, the shape is represented either using generative parametric models or explicitly through landmarks position. Parametric models have the advantage of yielding a compact shape representation featuring desired properties, such as shape smoothness and node-to-node correspondences (Shahim et al., 2013). In order to produce a coarse, but robust estimation of the bone anatomy, and based on our previous experience in (Chandran et al., 2015) for trabecular bone prediction from QCT images, we cast our solution to the shape regression as a multi-output extra-tree regression problem (Dumont and Marée, 2009; Marée et al., 2013), where an input feature vector $\boldsymbol{v}$ is mapped to an output response $\boldsymbol{w}$. Compared to previous registration-based approaches, such as deformable models, atlas-based, Active Shape Models, the motivation behind a regression-based approach is to enable a fast and optimization-free coarse shape estimation. In addition, as it is based on a compact parametric shape descriptor, the risk of overfitting is lessened considerably. Moreover, it is suitable for practical use as the parametric shape descriptors are derived from distinguishable femoral landmarks.

Table 1
Morphometric features used as input variables for coarse shape regression (Step 1). Fig. 1 offers for an illustration of the underlying internal coordinate system used to extract these features.

Distance	Min.Cross.Area	Angle
FHC-FNC	FNC Plane	CCD
FHC-MP	NEP Plane	
FNC-MP	SDP Plane	
FNC-SDP		
Femoral Head Diameter		

As input features, we adopt a set of morphometric features extracted from an internal coordinate system (Fig. 2) due to their robustness and simplicity of extraction (see Appendix A). Those features are subsequently computed on the QCT image. The redundancy of the features is evaluated via a Principal Component Analysis (PCA), leading to a final input feature vector $\boldsymbol{v} = (v_1, .., v_m) \in \mathbb{R}^m$ (m=9 in our experiments) as shown in Table 1.

To model the output response, each shape is represented by $\boldsymbol{w} \in \mathbb{R}^c$ parameters of a Statistical Shape Model (SSM) including the inner and outer meshes by projecting the shape onto the SSM space with c principal components, leading to an output response vector $\boldsymbol{w} = (w_1, .., w_c),\ w_i = \alpha_i\sqrt{d_i}$ (see Section 2.1.1 for details of the SSM model). We remark that for our purpose (shape regression), the compactness of this type of model is suitable for supervised learning where the number of samples is limited in comparison to the dimensionality of each sample (curse of dimensionality). Consequently, given a training set $\boldsymbol{SR} = \{\langle \boldsymbol{v}^{(i)}, \boldsymbol{w}^{(i)} \rangle | i = 1, ..., f|\}$ for the multi-output regression problem, we follow the hypothesis h_1: $\boldsymbol{v} \mapsto \boldsymbol{w}$ that maps from a space of features $\boldsymbol{v}$ to the space of responses $\boldsymbol{w}$.

The multi-output extra-tree algorithm builds an ensemble of T regression trees and requires no bootstrap sampling (Geurts et al., 2006; Marée et al., 2013; Dumont and Marée, 2009). To split the nodes of the tree, K alternate input features v_1, ..., v_K among the m number of features are randomly selected. In order to choose the best set of features for each K alternate input, a score based on Eq. 1 is calculated. The score maximising the variance reduction (maximum score) is then chosen to split the node. Considering D as the set of samples at the current splitting node, the score measure is defined as:

$$Score(c_i, D) = \frac{var(\boldsymbol{w}|D) - \frac{|D_l|}{|D|} var(\boldsymbol{w}|D_l) - \frac{|D_r|}{|D|} var(\boldsymbol{w}|D_r)}{var(\boldsymbol{w}|D)}. \quad (1)$$

$$var(\boldsymbol{w}|D) = \frac{1}{|D|} \sum_{i=1}^{|D|} \left\| \boldsymbol{w}^i - \frac{1}{|D|} \sum_{i=1}^{|D|} \boldsymbol{w}^i \right\|^2,$$

where c_i corresponds to the cut-point, D_l and D_r denote the left and right samples of the split node. For each node, the cut-point c_i value is uniformly drawn from the range $[v^D_{min}, v^D_{max}]$. Each node of the tree is recursively split until it reaches the minimum sample size n_{min}. Once the tree is grown, the leaf is assigned with a vector $\hat{\boldsymbol{w}}_L$, which is obtained as the average of the target vector values $\boldsymbol{w}^{(i)}$ associated to the samples falling in that leaf D_L (Eq. 2). The estimates of each tree is then aggregated by arithmetic average to

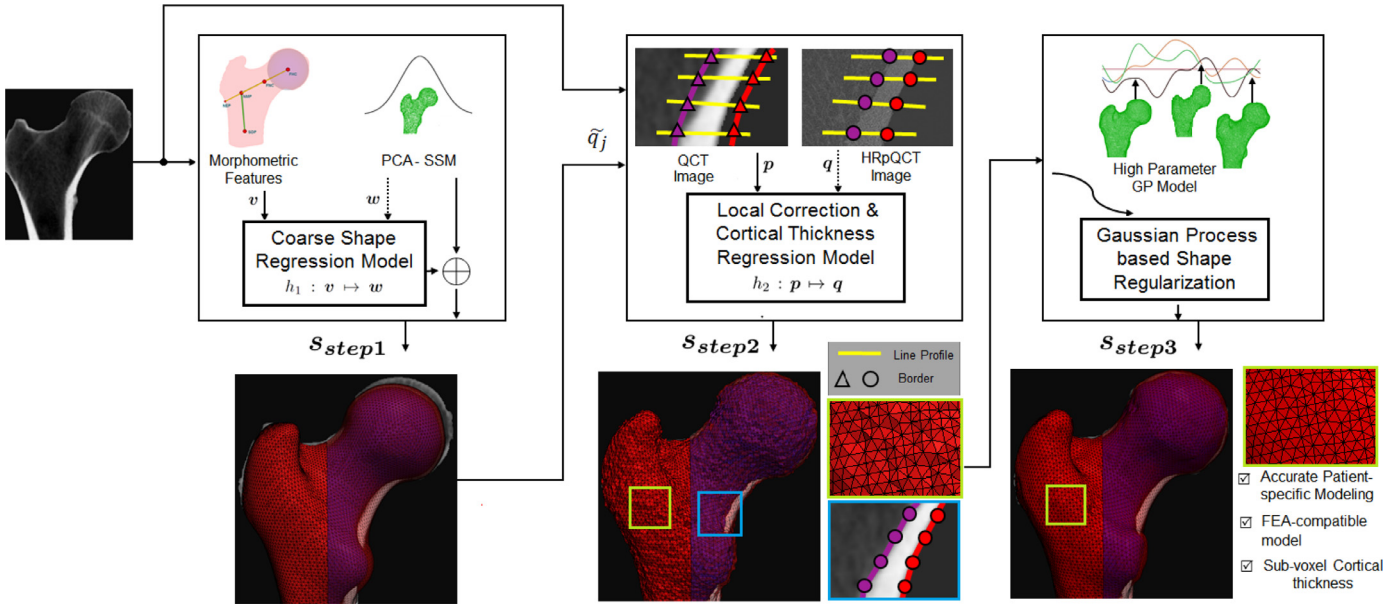

Fig. 1. Overview of the three steps that compose the proposed approach. First (Step 1), a coarse shape regression is performed via supervised learning based on morphometric features and parameters of a statistical shape model (SSM) of the proximal femur. This step yields two surface meshes that only approximate the trabecular and cortical bone regions. Those surface meshes are composed of triangle elements connected via nodes. Then (Step 2), local corrections are performed on the nodal positions via a supervised learning based on image features extracted from the QCT and corresponding HRpQCT images during the training phase. The corrected meshes now present a rough aspect but fit the bone regions, enabling accurate estimation of cortical thickness. In Step 3, the meshes are finally regularised after projecting the corrected meshes onto the space of a Gaussian process model featuring a two-kernel model, enabling smooth and shape-awareness priors. To highlight the improvement between the steps, the surface meshes (red: outer cortical surface, purple: inner cortical surface as the interface with the trabecular bone) are superimposed on the QCT image.

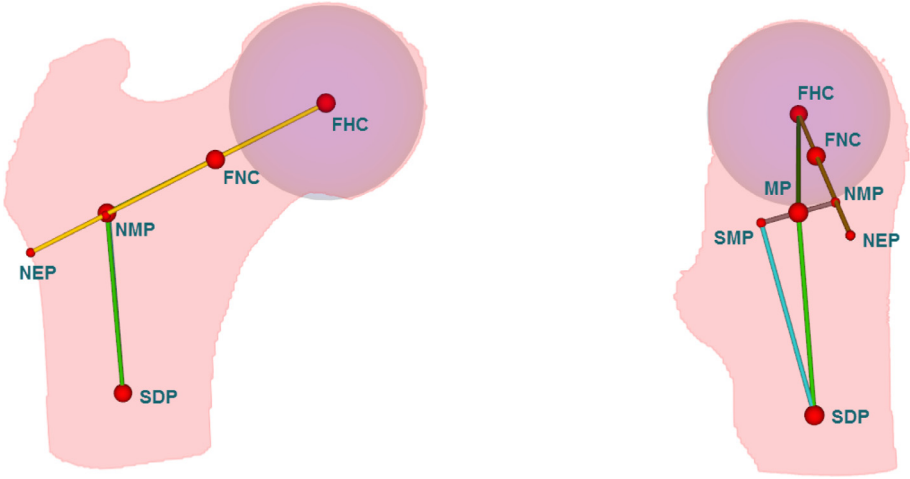

Fig. 2. Implicit coordinate system of the human proximal femur. Left : Coronal view. Right: Sagittal view. Acronyms: FHC - Femoral Head Center, FNC - Femoral Neck Center, NMP - Neck-axis Mid Point, NEP - Neck-axis End Point, SMP - Shaft-axis Mid Point, MP - Mid Point, SDP - Shaft-axis End Point. Adapted from (Chandran et al., 2015; 2017).

yield a final prediction:

$$\hat{\boldsymbol{w}}_L = \frac{1}{|D_L|} \sum_{i=1}^{|D_L|} \boldsymbol{w}^i. \tag{2}$$

For any new test image, an internal coordinate system is first constructed and the mean shape is aligned to it, as the same coordinate system has been predefined on it. This is a strong feature as no registration, involving optimization, is therefore necessary (we hence refer to this as "registration-free"). Morphometric features are then extracted (Table 1) and the shape regression is performed with the predicted model parameters $h_1(\boldsymbol{v}^{(test)})$: $\boldsymbol{v}^{(test)} \mapsto \boldsymbol{w}^{(test)}$. The resulting surface mesh from this step can then be obtained as $\boldsymbol{s_{step1}} = \boldsymbol{\mu} + \sum_{i=1}^{c} w_i^{(test)} \boldsymbol{u_i}$, and yields a coarse shape estimation, that can be regarded as an initialisation, since local regions of the surface mesh $\boldsymbol{s_{step1}}$ do not accurately fit the inner and outer cortical surfaces, hence the further need for a step where local corrections and thickness estimation take place.

2.1.1. Offline step: PCA-based Statistical Shape Model

In this section we provide details on the construction of the SSM model used in step 1. For shape representation in step 1, we adopted a PCA-based Statistical Shape Model (SSM) (Cootes et al., 1995). Initially, inner surface meshes were constructed from segmented inner masks of the HRpCT images. Then, for each, the corresponding outer surface mesh nodes were found by projecting line profile in the normal direction to the inner surface and using the segmented outer mask of the HRpQCT image. When needed manual corrections were performed to prevent the line profiles from crossing. This resulted in two-layered surface meshes representing the inner and outer cortex. Finally, all the two-layered

surface meshes of the femoral population $\{\Gamma^{(1)}, ..., \Gamma^{(n)}\}$ were aligned to a common coordinate system based on the femur's morphology as shown in Fig. 2. The Frobenius norm of the stretch tensor of the local gradient was selected to quantify the distance between the femurs of the population (Chandran et al., 2017). Based on those distances, the closest femur to the Frechet mean of the population was chosen as the reference template mesh $\Gamma^{(R)}$. Each shape $\Gamma^{(i)}$ was then represented by a discrete set of nodes, i.e.

$$\Gamma^{(i)} = \{x_k^i | x_k \in \mathbb{R}^{3N}, k = 1, ...N\} \tag{3}$$

where N denotes the number of nodes. Dense node-to-node correspondence between the surface meshes are established based on non-rigid shape registration. The Gaussian process registration framework (Lüthi et al., 2017), implemented in Scalismo (Scalismo, 2016) was used. Then, the coordinates for all k nodes are concatenated to one vector $\boldsymbol{s}$ that describes the shape:

$$\boldsymbol{s}^{(i)} = \{x_{1x}^i, x_{1y}^i, x_{1z}^i, ..., x_{Nx}^i, x_{Ny}^i, x_{Nz}^i\} \tag{4}$$

The mean shape $\boldsymbol{s}$ and its corresponding covariance matrix $\boldsymbol{S}$ is given by:

$$\boldsymbol{\mu}(=)\bar{\boldsymbol{s}} := \frac{1}{n-1}\sum_{i=1}^{n}\boldsymbol{s}^{(i)}, \quad \boldsymbol{\Sigma}=\boldsymbol{S} := \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{s}^{(i)}-\bar{\boldsymbol{s}})(\boldsymbol{s}^{(i)}-\bar{\boldsymbol{s}})^T \tag{5}$$

Eigen decomposition of $\boldsymbol{\Sigma}$ yields $max(n-1, N)$ eigenvectors $\boldsymbol{u_i}$ (principal modes of variation) and eigenvalues $\boldsymbol{d_i}$ (strength of the variations). Shapes from the PCA model can then be drawn as

$$\boldsymbol{s} = \boldsymbol{\mu} + \sum_{i=1}^{n}\alpha_i\sqrt{d_i}\boldsymbol{u_i} \tag{6}$$

where $\alpha \in \mathbb{R}^{n-1}$ represent the model parameters.

2.2. Step 2: Local outer and inner surface corrections and cortical thickness estimation

To accurately estimate the outer and inner cortical surfaces, a super-resolution strategy based on supervised learning is proposed. Line intensity profiles are established between corresponding nodes of the initial outer ($\boldsymbol{s_{step1}}$, $\tilde{q}_{outer} \in \mathbb{R}^3$) and inner ($\tilde{q}_{inner} \in \mathbb{R}^3$) meshes (Fig. 3c). The length of each line profile is represented by the parameter l and centred on $(\tilde{q}_{outer} + \tilde{q}_{inner})/2$. Inspired by super-resolution techniques (Alexander et al., 2014; Lu et al., 2017), we resample the line profile to match the voxel size of the HRpQCT scan (0.082 mm) via a linear interpolation[1], generating L samples per line profile. On each sample, four features are extracted, i.e intensity values along with texture features such as energy, entropy, and Haralick coefficients with two neighbourhoods (Haralick and Shanmugam, 1973). We used backward sequential schemes for selecting these three texture features from a pool of texture features (Bouatmane et al., 2007). This resulted in a feature vector $\boldsymbol{p}_j \in \mathbb{R}^{4\times L}$. Similarly, the output response variables corresponding to the locations of the outer and inner cortical interfaces are extracted via the segmented mask from the corresponding HRpQCT images along each line profile (Fig. 3b) and are modelled as a vector $\boldsymbol{q} = [q_{inner}, q_{outer}] \in \mathbb{R}^2$.

Given a training set $\boldsymbol{PQ} = \{\langle \boldsymbol{p}_j{}^{(i)}, \boldsymbol{q}_j{}^{(i)} \rangle | i = 1, .., f, j = 1, .., F|\}$, we cast the local correction as a multi-output regression problem and seek an hypothesis h_2: $\boldsymbol{p} \mapsto \boldsymbol{q}$. For a new test line profile, the predicted response is $h_2(\boldsymbol{p}_{(test)})$: $\boldsymbol{p}_{(test)} \mapsto \boldsymbol{q}_{(test)}$. Local corrections are then defined by updating each nodal position $\tilde{q}_{outer} \mapsto q_{outer}$ and $\tilde{q}_{inner} \mapsto q_{inner}$ and the cortical thickness is finally computed as $\|\boldsymbol{q}\|$. In order to maintain consistency between outer and inner corrected surfaces, the minimum cortical thickness is set to 0.33 mm, which is the minimal observable bone thickness (Treece and Gee, 2015). The multi-output regression provides a local response at every node of the surface mesh, which results in a locally corrected shape $\boldsymbol{s_{step2}}$, and a highly accurate estimation of the cortical thickness.

From the two corrected surface meshes, A number of tools (e.g. GMSH (v2.12.0) (Geuzaine and Remacle, 2009)) can be used to produce tetrahedral meshes for FE analyses. However, since the refinements occurred independently for every node, the corrected surfaces are not anymore smooth and may even self-intersect. This either makes the generation of tetrahedral meshes impossible or results in distorted tetrahedral elements, especially in areas with thin cortex or high curvature (e.g. trochanteric fossa). Distorted elements would be a cause of poor convergence for any subsequent finite element analyses. A regularisation is thus conducted in Step 3 to produce high quality surface meshes while preserving an accurate patient-specific shape.

2.3. Step 3: Mesh regularization

To create smooth, yet accurate meshes without having to remesh, we employ a Gaussian process model (Lüthi et al., 2013) featuring a larger amount of parameters compared to the rather compact PCA-based model used in Step 1. The large number of parameters enables highly-detailed shape descriptions, while maintaining the desired smooth shape. However, unlike in Step 1, the Gaussian process model is now employed as a fine and flexible shape regularisation that features a smooth and a shape-aware prior. This lessens the curse of dimensionality or Hugues phenomenon occurring for a fixed number of samples.

The Gaussian process model is described by $GP(\boldsymbol{\mu}, k)$ with $\boldsymbol{\mu}$ the mean deformation and a kernel function $k(x, x')$ describing the desired smoothness features of the deformations. An interesting feature of the Gaussian process modelling is that kernels can be combined in a straightforward manner. We thus combined a PCA-based SSM kernel k_{SSM} with a Gaussian kernel k to generate a flexible model $GP(\boldsymbol{\mu}_{SSM} + \boldsymbol{\mu}, k_{SSM} + k)$. We approximated the continuously defined Gaussian process model to obtain a statistical shape model augmented with deformations, which are defined by the Gaussian kernel. To model additional flexibility, we combine the statistical shape model with additional non-informative deformations. In the Gaussian process framework (Lüthi et al., 2017) the type of deformation can be defined using a kernel and include a large class of deformations, such as splines and radial basis functions. The Gaussian process is modeled continuously and is then low-rank approximated to its principal components to receive a parametric representation. The parameter C defines the amount of components taken for the approximation. In principle, flexible models defined for fine local deformations need more parameters (large C) because of a higher degree of freedom. Smooth models without local deformations can be approximated with a lower amount of basis functions (small C). For instance, a large C parameter yields a strong smooth shape prior and less dependency on the femoral shape prior, whereas a small C parameter lessens the smoothness prior and favors the shape-awareness prior brought by the SSM kernel. The final surface mesh $\boldsymbol{s_{step3}}$ is also expected to have a higher mesh quality (element aspect ratio and distortion) as compared to the locally corrected surface $\boldsymbol{s_{step2}}$ because of local shape adapted smoothness. We remark that the estimated thickness $\|\boldsymbol{q}\|$ does not need to be re-estimated as in practice its overall variation is negligible. The surface meshes resulting from Step 3 can now be used directly for generating tetrahedral meshes for finite element purposes. While this is beyond the scope of this study, an example is given in Figs. 11 and 12.

[1] no benefits were found when using higher order interpolation

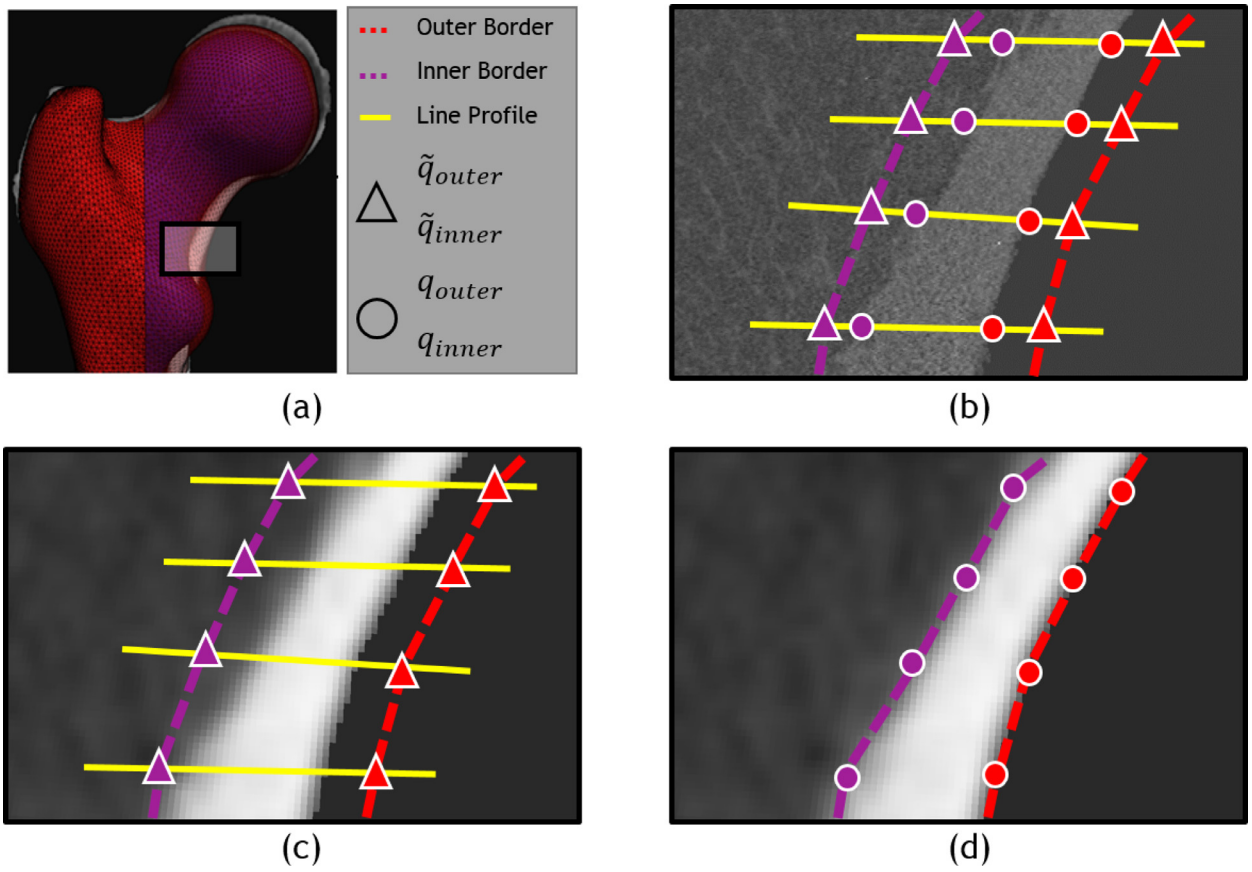

Fig. 3. Close-up on the local corrections performed in Step 2. a) The meshes produced in Step 1 for the inner and outer cortical surface are superimposed to the QCT image. The nodes clearly do no fit the QCT data and their position must be corrected. A local region of interest is highlighted in black b) The local corrections (i.e.discrepancies between triangle and circle symbols) are learnt from the HRpQCT image. First, line intensity profiles are defined between the nodes of the QCT-based meshes (red and purple triangles). Then, the nodes are moved along the line profiles at a subvoxel resolution (not shown in the figure) so as to match the cortex seen on the HRpQCT image and their position is updated (red and purple circles). (c) The original nodal position obtained on the QCT images is then corrected accordingly d) to mimic the cortical thickness learnt from the HRpQCT profile.

3. Experimental setup and results

3.1. Preparation of the image data

The study was performed on a set of QCT and HRpQCT images of human proximal femurs obtained from previous studies (Dall'Ara et al., 2013; Luisier et al., 2014). Briefly, 36 pairs (17 males, 19 females, 76 ± 12 years, range 46-96) were obtained from donors after ethical approval by the Medical University of Vienna. Each femur was scanned with a calibration phantom (BDC Phantom, QMR Gmbh, Germany) in a clinical QCT (Brillance64, Phillips, Germany, intensity: 100 *mA*, voltage: 120 *kV*, voxel size: $0.33 \times 0.33 \times 1.00$ mm^3) and HRpQCT (Xtreme CT, Scanco, Switzerland, intensity: 900 μA, voltage: 60 *kVp*, voxel size: $0.082 \times 0.082 \times 0.082$ mm^3). More information about the data are provided in Appendix C.4.

The HRpQCT images were automatically segmented by thresholding and two masks were generated for the cortical and trabecular regions using the "fill" algorithm described in (Pahr and Zysset, 2009) and manual corrections when necessary. A first surface mesh was created from the inner trabecular mask using the marching cube algorithm (Lorensen and Cline, 1987) and was resampled to a uniform triangle size and shape via Meshlab (v1.3.3). The outer surface mesh was then created by displacing all the nodes of the inner mesh along the normal of the corresponding triangle surface to fit the outer cortical mask. The outer surface mesh was then smoothed (Taubin, 1995; Belyaev and Ohtake, 2003). The average element size used for the present study was around 1 mm side length. The clinical QCT images were aligned to the corresponding HRpQCT images based on the internal coordinate system (Fig. 2) and rigid registration (Klein et al., 2010) when necessary. As a result, both QCT and HRpQCT images as well as the surface meshes are all aligned on the space of the HRpQCT images.

3.2. Model hyper-parameters

All parameters used in the present study were found heuristically through a grid-search on the parameter space (Table 2). For the SSM, the total number of nodes for $\Gamma^{(R)}$ is ${}^1N = 18100$ (9050 nodes for both outer and inner surfaces). The right femurs of the dataset were flipped resulting in 72 left femurs in total. Hence for each LOO study, an SSM was constructed from 70 femurs (leaving out the corresponding contralateral shape as well). For the shape regression (Step 1), the total number of samples for training ${}^1f$ was set to 70. The three hyper-parameters characterising the shape regression model were: ${}^1K = 9$ (the number of input features), ${}^1n_{min} = 2$ (minimum sample size) and ${}^1T = 40$ (number of trees). For the correction of the nodal positions (Step 2), the length of the line profile ${}^2l$ was set to 13.6mm. This results in a total number of (sub-voxel resolution) samples per line ${}^2L$ equal to 165 (13.6mm/0.082mm). The total number of samples for the training ${}^2f$ was set to 70 and ${}^2F$ was fixed to 9050 (nodes having correspondence between outer and inner surfaces). The three hyper-parameters characterising the correction model were ${}^2K =$

Table 2
List of the hyper-parameters and heuristically computed values used in the present study.

	Hyper-parameter	Value	Description
Step 1	${}^1K$	9	Number of input features
	${}^1n_{min}$	2	Minimum sample size
	${}^1T$	40	Number of trees
Step 2	${}^2L$	165	Number of sub-voxel resolution samples per line profile
	${}^2K$	660	Number of input features
	${}^2n_{min}$	5	Minimum sample size
	${}^2T$	80	Number of trees
Step 3	${}^3C$	400	Number of Gaussian model parameters

Table 3
Statistical description of the morphometric features of the femur population

Morphometric Features	Mean ± std	Range
Distance FHC-FNC(mm)	26.7 ± 3.8	17.0-29.9
Distance FHC-MP(mm)	47.7 ± 5.9	40.6-66.4
Distance FNC-MP(mm)	25.6 ± 5.5	21.4-40.6
Distance FNC-SDP(mm)	67.6 ± 6.4	52.8-80.3
Femoral Neck Min.Cross.Area(cm^2)	8.9 ± 1.8	6.3-12.9
NEP Min.Cross.Area(cm^2)	17.8 ± 3.3	11.5-22.8
SDP Min.Cross.Area(cm^2)	8.3 ± 2.0	4.3-13.8
Femoral Head Diameter(mm)	46.4 ± 4.0	41.5-54.4
CCD(°)	128 ± 7	112.5-143.1

660 (number of input features), ${}^2n_{min} = 5$ (minimum sample size) and ${}^2T = 80$ (number of trees). For the regularisation procedure (Step 3), the number of ${}^3C$ parameters for the Gaussian process deformation model was set to 400. Several sensitivity analyses regarding hyper-parameters were performed, and explained below for each step of the proposed approach.

3.3. Coarse shape regression (Step 1)

Given the variability of the morphological features, as measured in our dataset (see Table 3), it was important to ensure that the coarse shape regression (Step 1) was robust enough to enable meaningful local corrections (Step 2). We therefore performed three different sensitivity analyses as detailed below.

For the coarse shape regression, root mean square error (RMSE) and Hausdorff distance (HD) were evaluated to determine the influence of the number of principal components of the SSM on the accuracy and the robustness of the outer mesh prediction. The set of morphometric features was kept constant during the procedure (Table 1). Based on a LOO strategy, we observed that increasing the number of modes did not significantly influence either RMSE or HD (paired Wilcoxon test, $p > 0.5$). Hence, the number of modes for the rest of the study was thus set to 5 in order to have a compact model. Conversely, a severe reduction on the number of modes below 4 produced a negative influence on the RMSE and HD metrics.

On a second analysis, feature importance was evaluated based on the variance measure of the multi-output extra tree regressor (Geurts et al., 2006). The results are presented in Fig. 4. The femoral head diameter appears as the most important feature followed by the CCD angle.

A third sensitivity analysis was performed to evaluate the impact of possible errors introduced during the extraction of morphometric features (Step 1) on the RMSE and HD metrics. We adopted a Monte-Carlo uniform sampling strategy where a large number of 70 perturbations were sampled assuming a Gaussian distribution with standard deviation (SD) of each feature (Table 3). We sampled perturbations up to 2*SD, which we think sufficiently models potential and realistic deviations (given by sample population statistics, as opposed to random) on the extraction of morphological features that are used by the system to infer shape. Based on a LOO strategy, a multi-output regression was performed and the

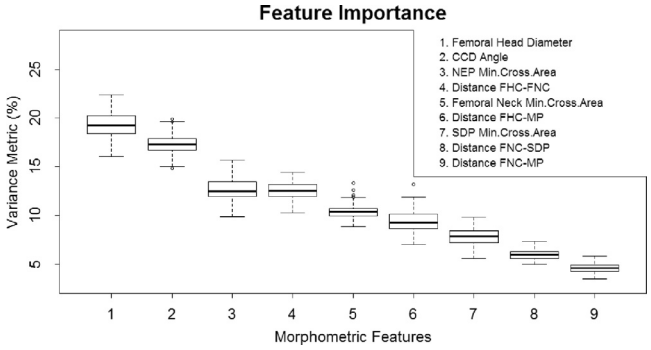

Fig. 4. Respective importance of the morphometric features used in Step 1.

accuracy and robustness of the shape predictions were again evaluated in terms of RMSE (Fig. 5) and HD (Fig. 6). It can be noticed that RMSE and HD errors were only sensitive to the femoral head diameter.

3.4. Cortical thickness estimation (Step 2)

The error in the cortical thickness estimation produced after Step 2 was defined as $\|\boldsymbol{q_{GT}}\| - \|\boldsymbol{q}\|$, with $\|\boldsymbol{q_{GT}}\|$ the thickness measured on the HRpQCT images (ground-truth) and $\|\boldsymbol{q}\|$ the thickness estimated in Step 2. Both the thickness and its deviation from the ground-truth were evaluated based on a LOO strategy. Following (Treece and Gee, 2015), we evaluated the ability of the proposed approach to estimate the cortical thickness with sub-voxel accuracy for three thickness ranges (Fig. 7). A fairly good thickness estimation accuracy was observed for thick cortical areas (between $3mm$ and $6mm$ thick) with a mean ± std error equal to $-0.01 \pm 0.38\,mm$. A slight over-estimation was measured when the cortex is the thinnest with a $0.11 \pm 0.45\,mm$ error for the $0.33mm - 1mm$ range. Thickness errors were $-0.01 \pm 0.32\,mm$ when thickness was between $1mm$ and $3mm$ and $0.05 \pm 0.40\,mm$ overall.

The 25th, 50th and 75th percentiles of the error in cortical thickness estimation were calculated for each node of the mean femur to illustrate their spatial distribution (Fig. 8). In addition, the average "ground-truth" thickness obtained from HRpQCT was also calculated for each node to represent the spatial variation of the cortical thickness for the femoral population used in the study (Fig. 8a). Qualitative results reveal that estimating cortical thickness is more challenging in the femoral head region (low thickness range).

3.5. Regularisation of the surface meshes (Step 3)

For each step, we evaluated the mean aspect ratio of triangle elements composing the meshes (the ratio of the longest and shortest edge of the element), their accuracy (RMSE) and their robustness (HD) to assess the improvements brought by Step 3. Fig. 9b

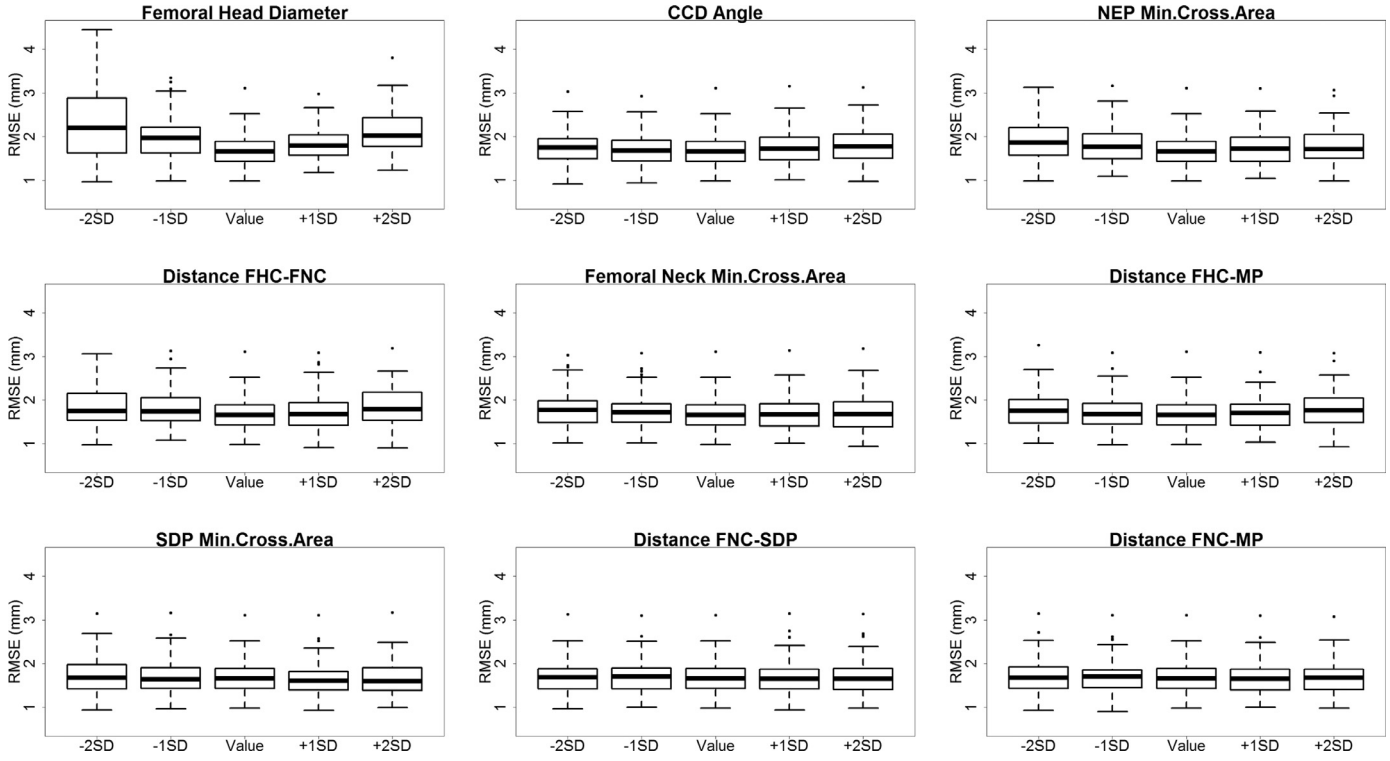

Fig. 5. Accuracy (RMSE-Root Mean Square Error) of the morphometric feature extraction performed in Step 1. Each boxplot corresponds to leave-one-out results on the studied population. Each feature was perturbed with an error corresponding to its standard deviation (SD)

confirms the expected result that the meshes resulting from Step 1 are the least accurate and least robust (RMSE=$1.70 \pm 0.41mm$; HD=$5.97 \pm 1.79mm$). The aspect ratio of the elements generated by Step 1 is very close to 1 (AS=1.23 ± 0.01). An aspect ratio of 1 corresponds to a triangle element that is almost equilateral and less prone to distortion if used in FE analyses. Step 2 greatly improved both precision metrics (RMSE=$0.37 \pm 0.12mm$; HD=$3.61 \pm 1.62mm$) to the detriment of the aspect ratio (AS=1.97 ± 0.07). After Step 3, mesh accuracy, robustness and quality are improved (RMSE=$0.36 \pm 0.29mm$; HD=$3.90 \pm 1.57mm$, AS=1.40 ± 0.02).

4. Discussion

The objective of this work was to introduce an approach for generating automatically and without segmentation bone meshes from QCT scans for finite element analyses. The proposed approach consists of three steps: (i) shape regression based on morphometric features to get a rough estimation of the surface meshes, (ii) local corrections of the nodal positions based on supervised learning to fit at sub-voxel level the surface meshes to the bone surfaces and improve the thickness information, and (iii) shape regularisation based on a Gaussian process framework to improve the quality of the surface meshes. While being out of the scope of this work, we remark that using available tools such as GMSH (Geuzaine and Remacle, 2009), reliable tetrahedral meshes accounting for the cortical shell, can be created for FE analysis and analysis of bone strength and fracture risk in patients. The proposed methodology was applied on a representative and heterogeneous database of 72 human proximal femurs scanned with both clinical QCT and HRpQCT modalities (Dall'Ara et al., 2013; Luisier et al., 2014).

4.1. General discussion

Step 1 yields a coarse shape prediction based on morphometric features extracted from the internal coordinate system of each proximal femur. This step has many advantages over other approaches based on shape atlases or statistical shape models. First, the extraction of the internal coordinate system is robust, fast, easy, and well-tailored to manual interventions in a clinical setup (Chandran et al., 2015). Secondly, the procedure is "registration-free" as no optimization based on a registration metric is required. This reduces the computational time, but also provides morphometric features that are independent from the position and orientation of the bone, unlike point-distribution models (Blanc et al., 2012). A thorough sensitivity analysis was conducted on the surface meshes resulting from Step 1. This is in our opinion very important as to ensure that the complete pipeline is robust and can be translated into practice. In this regard, our experiments on the number of modes to be retained for shape regression suggests that five principal components suffices to have a stable coarse shape regression. Including more modes of variation did not lead to any improvements, as higher modes describe small shape variations in shape. The quality of the predicted shape also depends on both the descriptive power of the features and the sensitivity of the model to perturbations on the extracted features. Our results demonstrate the importance of the femoral head size as a descriptive feature, along with a higher sensitivity of the model to any perturbations of this morphometric feature (Figs. 5, 6 top-left). Conversely, the model is less sensitive to perturbations on other morphometric features. These two aspects suggest that shape characterisation of the proximal femur is overall driven by the femoral head, however, we cannot conclude that this feature alone can be used in practice since our experiments were based on the complete set of described morphometric features (Table 1). Besides, a longer shaft than seen in our femurs might also be an important feature.

The local corrections step (Step 2) significantly improved the shape accuracy of the $\boldsymbol{s_{step1}}$ meshes at the cost of a higher element aspect ratio and a rough surface. These two aspects are critical if a volumetric mesh composed of tetrahedrons were to be

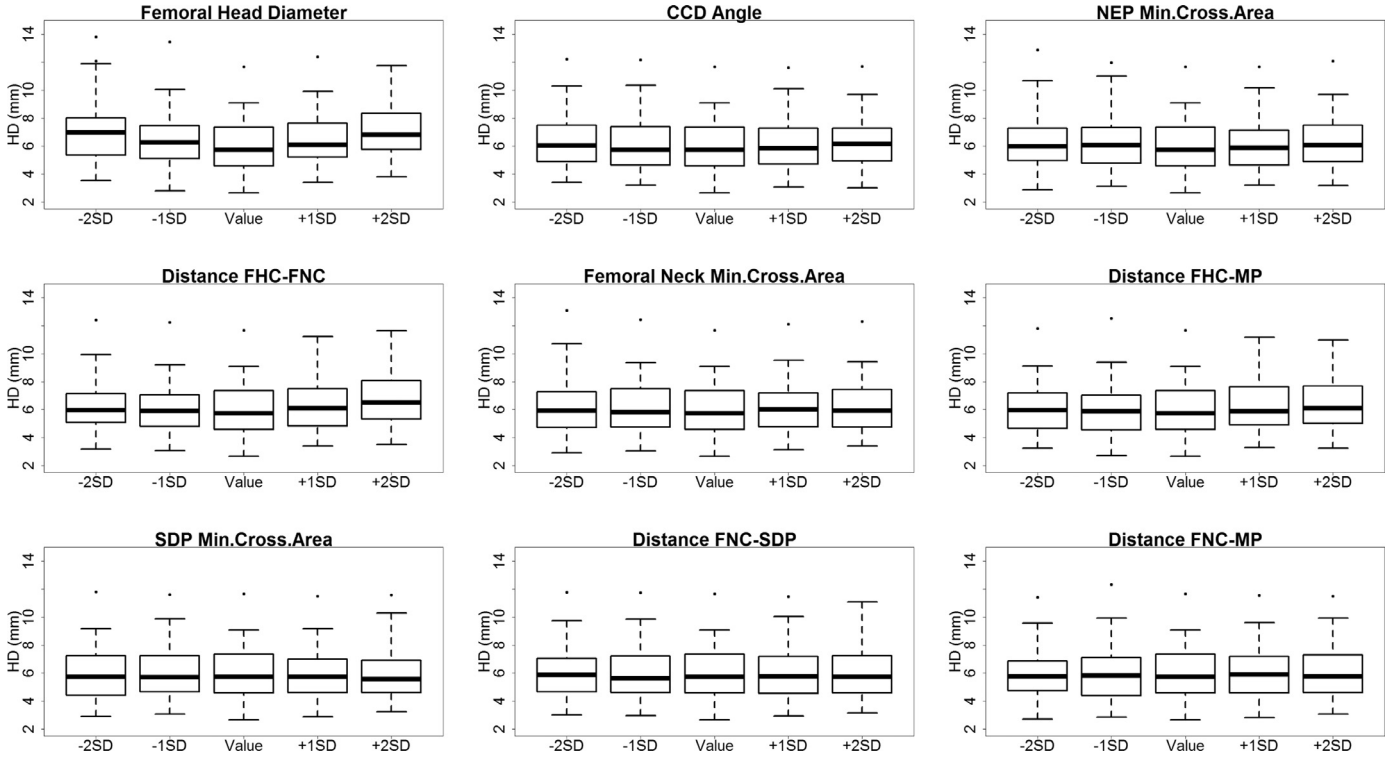

Fig. 6. Robustness (HD-Hausdorff distance) of the morphometric feature extraction performed in Step 1. Each boxplot corresponds to leave-one-out results on the studied population. Each feature was perturbed with an error function using the standard deviation (SD) of each feature on the studied population.

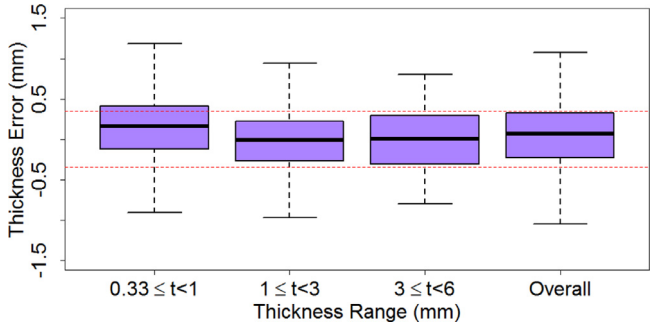

Fig. 7. The errors in the cortical bone thickness estimated from QCT for three thickness ranges. Dotted lines indicate sub-voxel accuracy. Positive and negative errors respectively correspond to under- and over- estimation.

generated from the two surfaces meshes for finite element analyses. Thanks to the sub-voxel sampling and modeling through the pairs of HrpQCT and QCT images, the proposed approach yields sub-voxel accurate cortical thickness estimations (Fig. 7), especially around the femoral neck region (Fig. 8). In comparison, larger errors were observed around the femoral head, which we remark to be less relevant than the neck area, being the main region at risk of fracture. The incorporation of a shape regularisation step based on a Gaussian process model in Step 3 combines the desired a priori information such as shape smoothness and statistical shape variability through kernels of different nature. Thanks to such flexibility, the resulting meshes can easily respect the local bone morphology (a small kernel can be used locally for areas of high curvature such as the trochanters, larger kernels for rather smooth regions such as the shaft). The regularisation process improves the quality of the meshes (Fig. 9c) while maintaining good levels of accuracy (Fig. 9a-b). It is also fast and does not need any optimisation.

In terms of overall computational efficiency, the machine learning training step of the proposed approach is the more time-consuming part, however this is performed in an offline manner and is only performed once. During testing, the proposed approach delivers a patient-specific model with sub-voxel accuracy of cortical thickness estimation within 3 minutes on a 3.2 GHz Intel Core i7 processor. Out of the three steps, step 2 (cortical thickness estimation) is the more time-consuming one, as it extracts imaging features for each line profile. Steps one and three are time-effective operations in the order of 2min and milliseconds, respectively.

4.2. Comparison with other techniques

In Treece and Gee (2015), the authors compared several techniques for cortical thickness estimation and reported the superior performance of their approach. As this study used the same dataset as our work, we compared our results to theirs and found that our respective approaches compare favourably (Treece and Gee, 2015). In particular, the thickness errors arising from our method were of $0.11 \pm 0.45\,mm$ for the $[0.33mm - 1mm]$ thickness range, $-0.01 \pm 0.32\,mm$ for the $[1mm$ - $3mm]$ range and $-0.01 \pm 0.38\,mm$ for the $[3mm$ - $6mm]$ range. On the very same dataset, Treece and Gee report thickness errors of $-0.15 \pm 0.23\,mm$, $0.12 \pm 0.39\,mm$ and $0.04 \pm 0.25\,mm$. This finding is very encouraging, especially considering that Treece and Gee's approach has already been used in clinical studies (Treece et al., 2015; Poole et al., 2017) and clinical trials (Whitmarsh et al., 2016). Nonetheless, it is worth mentioning that this comparison is limited as we remark that the study of Treece et al. (Treece et al., 2015) utilized a ground-truth generation based on full width at half maximum (FWHM), while in our study the HRpQCT images were semi-automatically segmented by thresholding with manual corrections, and two masks were generated for the cortical and trabecular regions using the "fill" algorithm described in (Pahr and Zysset, 2009).

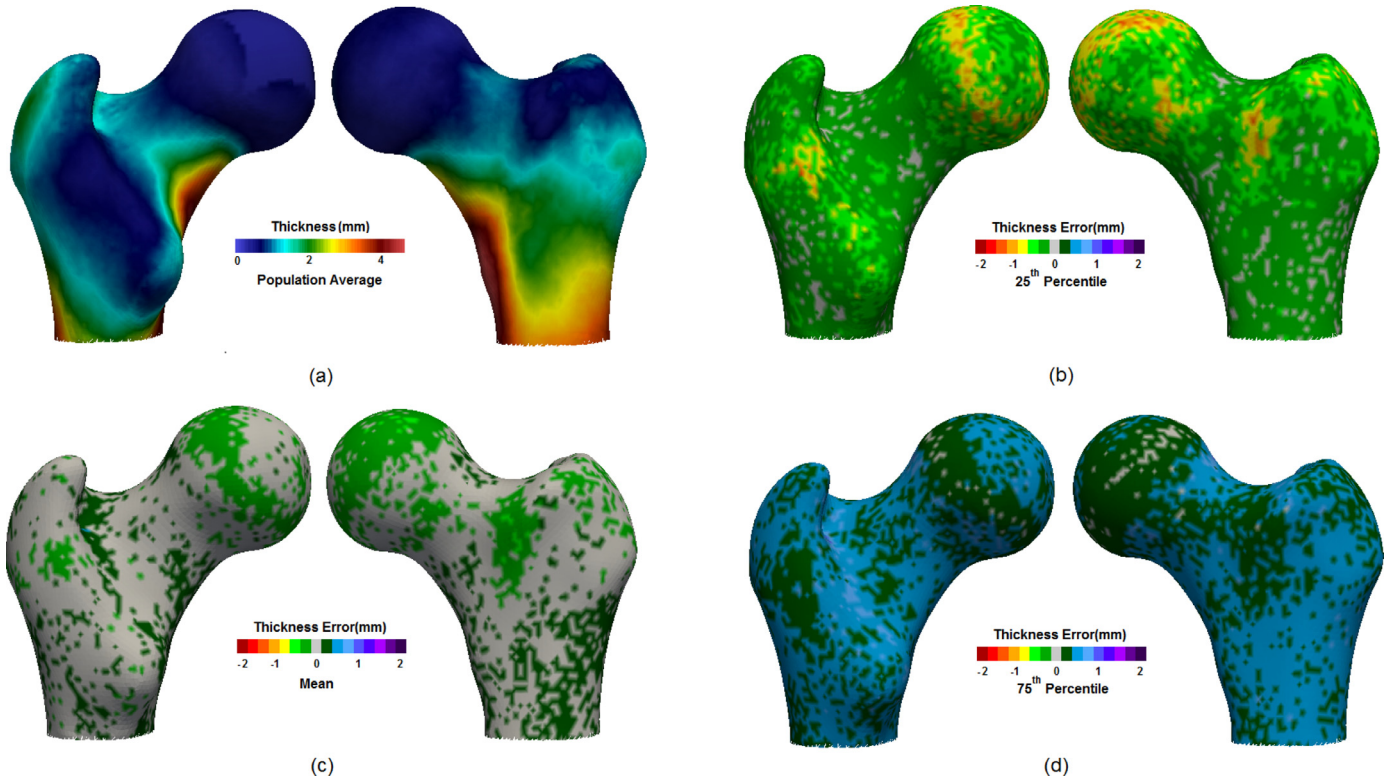

Fig. 8. Left and right views of the spatial distribution of the cortical thickness and its error on the reference template femur (see Section 2.1.1, $\Gamma^{(R)}$). (a) Average ground-truth thickness measured from HRpQCT images of the whole population. (b) 25^{th} percentile error, (c) 50^{th} percentile error, and (d) 75^{th} percentile error.

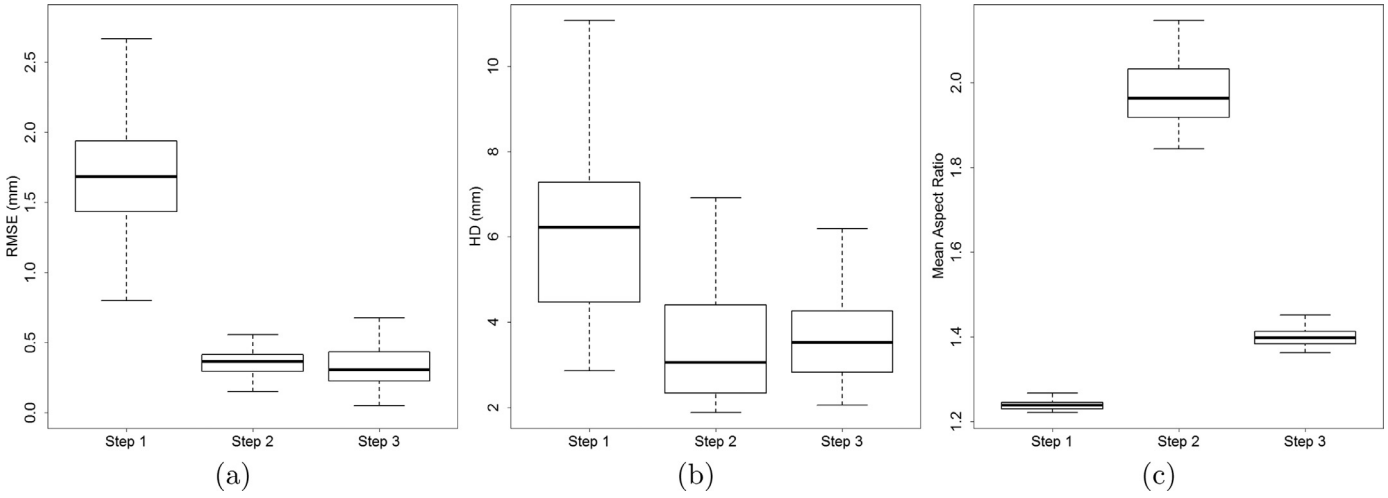

Fig. 9. Evaluation of the a) accuracy (Root Mean Square Error - RMSE), (b) robustness (Hausdorff distance - HD) and (c) element quality (Mean Aspect Ratio) for the surface generated by each step of the presented approach. Each boxplot corresponds to leave-one-out results.

From our previous experience and extensive evaluation of image-based registration approach involving rigid, affine and b-spline registration for proximal femoral image analysis (Chandran et al., 2017), we noticed that, for this specific task where the cortical thickness spans 0.3–3 mm, the results of the image registration are not accurate enough (average error superior to 3 mm) in determining cortical thickness of the femur. In particular, in our experience, obtaining accurate registration results for a bone cortical thickness of less than 1 mm, which is important for biomechanical simulations, becomes a major challenge. Henceforth, we selected a landmark-based approach since it was able to perform at sub-voxel accurate cortical thickness prediction while keeping a low algorithmic parametrization, which we esteem important for further clinical use of the proposed solution. Similarly, our experience showed that while an image registration approach could be used for extracting morphometric features, the error in the predicted morphometric features lies within the upper limits used for the sensitivity analysis.

Due to its statistical nature, the accuracy of our machine learning approach depends on the quality of the training database, which needs to be representative and of sufficient quality. For example, a predictive model built on another scanning protocol can introduce errors in the prediction of cortical thickness unless it is re-trained. The machine learning approach also requires pairs of clinical and high resolution images for the training. In comparison, most approaches are generative methods that work directly

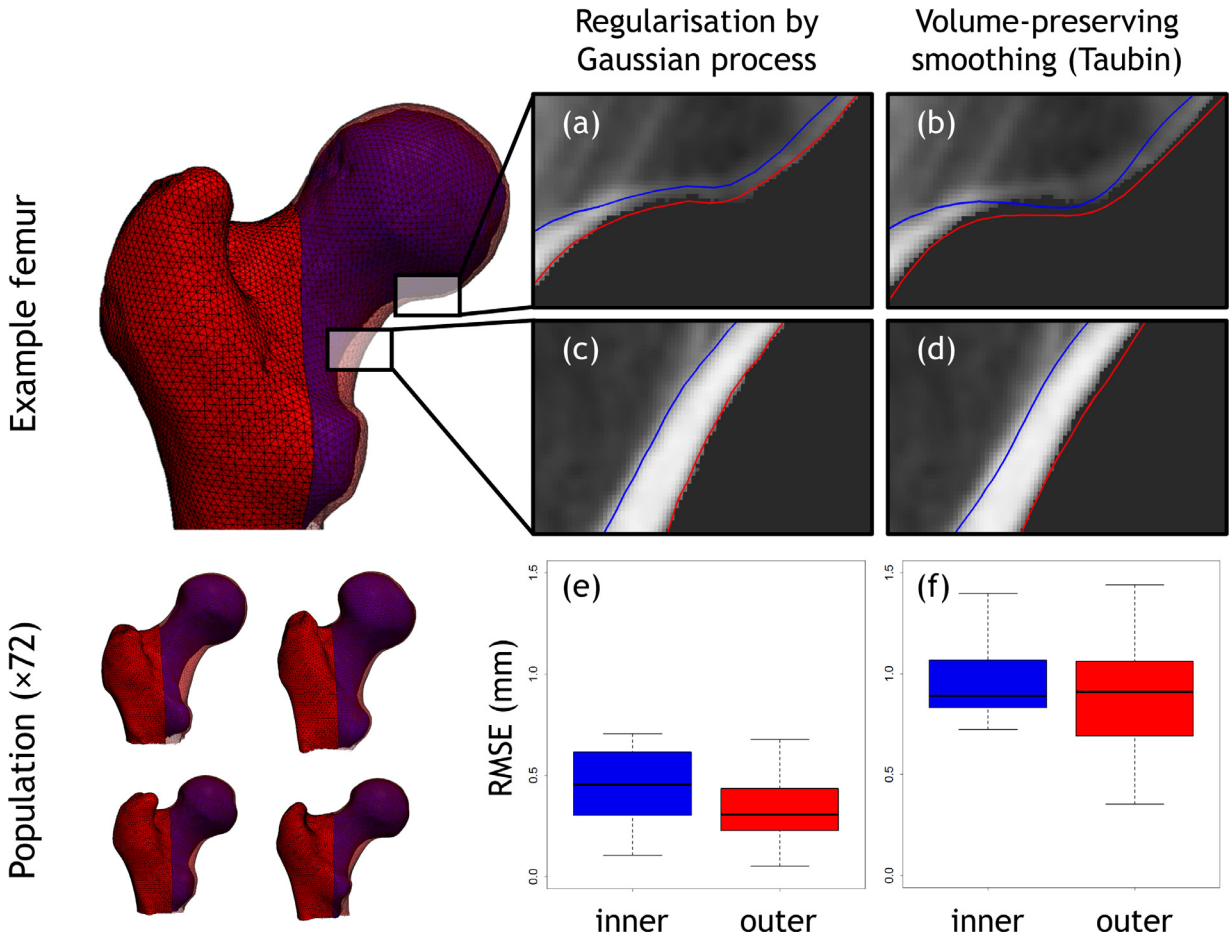

Fig. 10. Comparison of the accuracy (RMSE) of the inner and outer surfaces obtained after the regularisation by Gaussian process (Step 3) and a volume-preserving smoothing (Taubin algorithm). The parameters for the Taubin smoothing (λ, μ) were obtained heuristically so as to match the aspect ratio achieved with the regularisation (Fig. 9c). The visual inspection on the mean femur of our population shows that, unlike regularised surfaces (a), surfaces treated by Taubin smoothing (b) tend to deviate from the cortical borders with high curvature. However, both perform similarly in regions of low curvature (c, d). An analysis over the whole population shows that higher accuracy is achieved with the proposed regularisation (e) than with Taubin's smoothing (f).

on clinical images. Yet, they require nonetheless a fine-tuning of the model parameters to get an optimal solution (Pakdel et al., 2012), which must still be validated against a ground-truth provided by high resolution images (Treece and Gee, 2015; Humbert et al., 2016).

It can be argued that the upsampling of the line profile to HRpQCT voxel size in Step 2 might result in redundant data. However, along with the intensity values, the super-resolution strategy uses second-order texture features to capture the variability within the neighbourhood of line profile and provides sub-voxel cortical thickness accuracy similar to fine tuned generative methods (Lu et al., 2015; 2017).

It is worth mentioning again that the purpose of most previous works is mainly to measure cortical thickness (Pakdel et al., 2012; Humbert et al., 2016; Treece and Gee, 2015; Zhang et al., 2016). As presented here, we do that as well, but also aimed at generating surface meshes that can be used for producing tetrahedral meshes for FE analyses. Mesh generation from clinical images normally starts with a segmentation that often requires time-consuming manual corrections. As inaccuracies at that level would have consequences on a FE analysis (Peleg et al., 2014), some authors proposed to deconvolve the CT images so as to improve the segmentation (Falcinelli et al., 2016). Our study suggests that approaches based on machine learning are able to produce high-quality meshes without such pre-processing step.

Generative methods estimate cortical thickness along the line profile following the normals of the outer mesh. In principle, an inner mesh could be generated by shifting inward the outer nodes according to the measured thickness. Yet, the normals often cross each other in areas of high curvatures, leading locally to intersecting mesh elements. This prevents the generation of tetrahedral elements and FE analysis. We thus follow a different paradigm. Our inner and outer surface meshes have nodal correspondence (Step 1) and the pairs of nodes are moved until the line profile matches the cortical borders (Step 2), which maintains their topology.

The resulting surfaces from Step 2 are rough and must be regularised (Step 3). A number of volume-preserving approaches exists (e.g. Taubin's algorithm (Taubin, 1995)), but smoothing is then done globally and in our experience with these methods, they needs to be tuned for each bone. Ultimately, this deteriorates the surface accuracy and thickness estimation. On the other hand, the regularisation that we propose adapts to the topology of the inner and outer surface meshes simultaneously ("shape-aware"). The idea is to account for two types of priors, namely *"we want a smooth mesh"* and *"we want a correct bone shape"*, which can be obtained through the Gaussian process model proposed herein (Lüthi et al., 2013; Shahim et al., 2013; Lüthi et al., 2017). A comparison between our regularisation and a Taubin smoothing is available in Fig. 10.

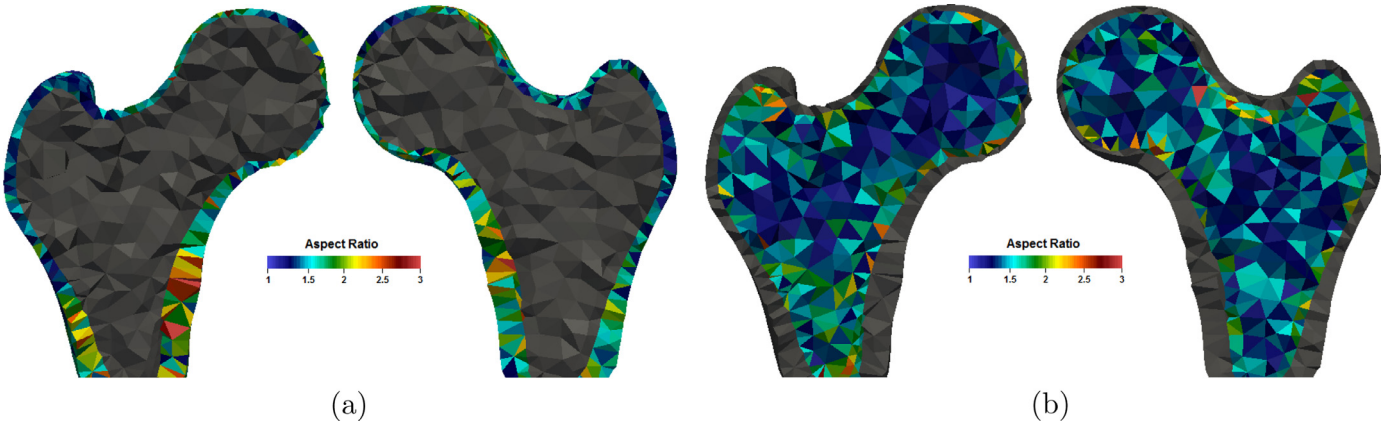

Fig. 11. Left and right views of the spatial distribution of the tetrahedral elements generated from a random pair of s_{step3} surface meshes. Their aspect ratio in the cortical (a) and trabecular (b) regions was evaluated. The measured value was 1.35 ± 0.05 in the cortex and 1.25 ± 0.03 in the trabecular bone. Those values correspond to high-quality elements and fall within the range used by Abaqus for FEA (1;10), 1 corresponding to the aspect ratio of a regular tetrahedron (best case) and 10 being the value for considering the element as degenerated (worst case).

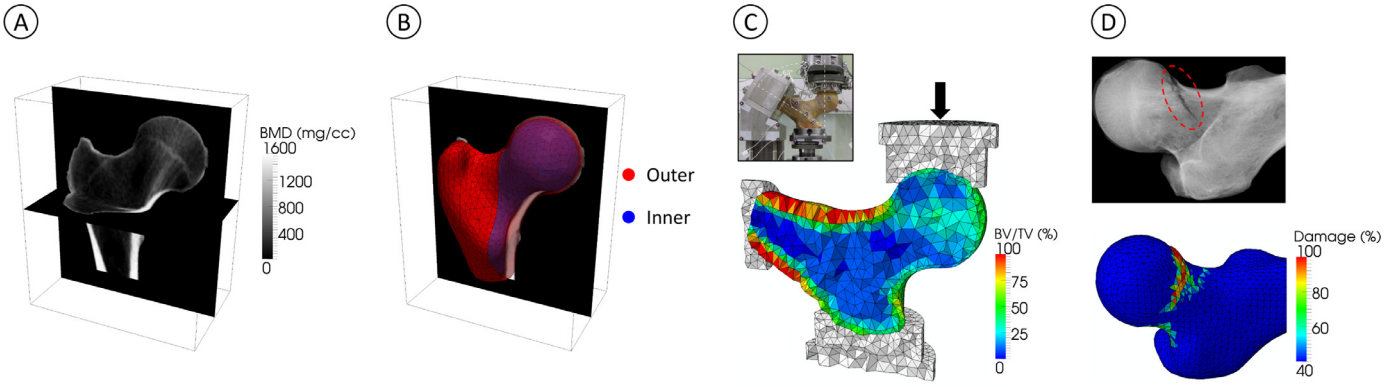

Fig. 12. A QCT scan (A) is used by our approach to generate surface meshes for the inner and outer cortical surfaces (B). (C) A tetrahedral mesh is built via GMSH, positioned and embedded (extra grey elements) as in the in vitro experiment. Axial loading is then conducted in Abaqus (v6.132, Simulia; Dassault Systemes, Velizy-Villacoublay, France), the mechanical behaviour the bony elements being simulated by a density (BV/TV)-based material model. (D) The simulated damage area matches the failure observed on the X-ray acquired after the experiment.

4.3. Summary of the strengths of the approach

Our goal is to integrate an accurate cortical thickness estimation in a mesh that can be used for finite element analyses. In regards of the proposed solution, four assets are highlighted.

First, the internal coordinate system for shape regression avoids the need for an actual registration during the initialisation step ("registration-free"). To form this coordinate system, few landmarks are needed (Chandran et al., 2017). We conducted a sensitivity analysis to know how precise they have to be (Figs. 5,6), indicating suitability of the approach for clinical use.

A second novelty is that the surface meshes produced in Step2 are "topology-preserving". That means that the outer (respectively inner) surface meshes generated for two individuals feature the same topology, no matter which bone is being analysed. The spatial comparisons of cortical thickness in large clinical studies can thus be more reliable as there is a direct correspondence between nodes/elements across the population.

The third strength of the pipeline lies in the modularity of the kernel design of the Gaussian process which allows combination of different kernels in a straightforward manner. This flexibility can be further explored by investigating other kernels or combinations thereof, which can be used to model pathological anatomies not known to a typical SSM-based kernel (Lüthi et al., 2013; 2017).

5. Conclusion and future work

Cortical bone thickness has been shown to contribute significantly to the fracture risk (Treece et al., 2015). Such information can be captured from clinical images, but is hardly included in QCT-based finite element models. Our approach produces topology-preserving finite element meshes with sub-voxel cortical thickness estimation from CT data. It is automatic and registration-free, and its computation time makes it suitable for integration into a system for clinical use.

In this study, the femur was chosen as an example of application, but this can be extended to other anatomical sites provided that high-resolution CT images are also available (e.g. the vertebra (Pahr et al., 2014), the wrist (Varga et al., 2009) or the kneecap (Latypova et al., 2016)).

One important limitation of the present study is the fact that it was conducted on in vitro images. This was necessary to build the pipeline, as HRpQCT scans cannot at present be acquired in vivo at the hip area. Future work includes evaluating its robustness (e.g scan-rescan) on in vivo CT data.

Finally, to provide some insights as to what can be done with this method, we produced a tetrahedral mesh from our surface meshes and evaluated its quality (Fig. 11). Since the femurs of our database have not only been scanned, but also tested mechanically until failure (Dall'Ara et al., 2013), we checked that this mesh was

suitable for finite element analyses and compared the experimental and simulated failure locations (Fig. 12). This pipeline will be expanded in a future work to determine how cortical thickness improves finite element predictions.

Acknowledgements

The authors would like to thank Enrico Dall'Ara (Vienna University of Technology, Austria) for sharing his QCT and HRpQCT images. We also gratefully acknowledge Graham Treece (University of Cambridge, UK) for his helpful comments as well as Dieter Pahr and Michael Kinzl (Vienna University of Technology, Austria) for providing the script generating the embedding on the finite element model. Finally, we thank Marcel Luethi (University of Basel, Switzerland) for his valuable inputs regarding the Gaussian process. This research was made possible through funding from the Swiss National Science Foundation (SNSF-143769) and the Gebert Ruef Foundation (GRS-079/14).

Appendix A. Construction of internal coordinate system

An internal coordinate system of the femur was constructed as shown in Fig. 2. First, the center of the femoral head (FHC) is defined by a mass center of a spherical region with maximal cross-section area. The neck axis is then computed by following the procedure reported by Kang et al. (Kang et al., 2005; 2003). In short, the radius of the spherical region of the femoral head is enlarged by 25%, and an initial neck center is defined. Using Powell's optimization (Press et al., 2007), the center of the femoral neck (FNC) is computed, and the neck axis is defined as the line between FHC and FNC (see Fig. 2). The intersection point between the neck axis and the lateral surface of the femur is defined as the neck-axis-end-point (NEP). Then, the mass center of slices distal to NEP are computed, followed by RANSAC fitting (Fischler and Bolles, 1981) to define the shaft axis. Generally, as the neck and shaft axes do not intersect, a midpoint (MP) is defined as the shortest distance between the neck and shaft axes. The most distal point of the shaft axis is chosen as shaft-axis-distal-point (SDP). By connecting SDP, MP and FHC, the implicit coordinate system is constructed. As morphological parameters we calculated known shape descriptors of the femur, such as the caput-collum-diaphyseal angle (CCD), femoral head diameter, and distances describing the femoral neck anatomy.

Appendix B. Offline step: Gaussian process model

For modeling the statistical shape models with and without additional flexibility we used the Gaussian process framework (Lüthi et al., 2017). In this approach a deformation model is defined on a reference domain, such as a surface or a volume, using a Gaussian process. More precisely let $\Gamma^{(R)} \subset \mathbb{R}^3$ be the reference surface and $u : \Gamma^{(R)} \to \mathbb{R}^d$ a deformation field to deform the reference surface, whereas a deformed shape is defined as

$$\Gamma_T = \{x + u(x) | x \in \Gamma^{(R)}. \tag{B.1}$$

We define a prior over the possible deformations using a Gaussian process $u \sim GP(\boldsymbol{\mu}, k)$, where $\boldsymbol{\mu} : \Gamma^{(R)} \to \mathbb{R}^3$ is a mean function and $k : \Gamma^{(R)} \times \Gamma^{(R)} \to \mathbb{R}^{3\times3}$ is a covariance function. The mean function defines the average deformation from the reference that we expect (which is typically the zero function, assuming that the reference is an average shape) and the covariance function defines the characteristics of the allowed deformations. The resulting model is a fully probabilistic model over femur shapes.

We can turn this infinite dimensional problem into a tractable one by approximating the model using a truncated Karhunen-Loève expansion. This leads to a parametric model $\tilde{u}(\alpha, x)$ of the form

$$\tilde{u}(\alpha, x) := \mu(x) + \sum_{i=1}^{C} \alpha_i \sqrt{\lambda_i} \phi_i(x),\ \alpha_i \sim \mathcal{N}(0, 1), \tag{B.2}$$

where $\lambda_i \in \mathbb{R}$ are weights and $\phi_i : \Gamma^{(R)} \to \mathbb{R}^3$ corresponding basis function. A detailed discussion about the accuracy of the low-rank approximation and the amount of basis-functions can be found in Lüthi et al. (2017).

B1. Sample covariance kernel (k_{ssm})

As reported in (Lüthi et al., 2017) that kernels can also be described from training data. A Gaussian process $GP(\mu_{ssm}, k_{ssm})$ that models these characteristic deformations is obtained by estimating the empirical mean

$$\mu_{ssm}(x) = \frac{1}{n} \sum_{i=1}^{n} u_i(x) \tag{B.3}$$

and covariance function

$$k_{ssm}(x, x') = \frac{1}{n} \sum_{i=1}^{n} (u_i(x) - \mu_{ssm}(x))(u_i(x') - \mu_{ssm}(x'))^T. \tag{B.4}$$

B2. Combining kernels

The deformation prior is defined using a Gaussian process $u \sim GP(\boldsymbol{\mu}, k)$ with a mean function $\boldsymbol{\mu}$ and a covariance function k. The approach allows to combine multiple kernels according to a simple set of rules (Lüthi et al., 2017). As one example, the sum of two kernels $h(x, x')$ and $g(x, x')$ gives a valid kernel again:

$$k(x, x') = g(x, x') + h(x, x') \tag{B.5}$$

As an example we combine $k(x, x')$ with $k_{SSM}(x, x')$ the following:

$$u \sim GP(\boldsymbol{\mu} + \boldsymbol{\mu}_{SSM}, k + k_{SSM}) \tag{B.6}$$

Appendix C. Data source: Additional information

The femurs used in this study also underwent dual-energy X-ray absorptiometry (Discovery QDR, Hologic Inc., USA) (Dall'Ara et al., 2016), which is the gold standard for screening osteoporosis. Values of aBMD (areal BMD) were thus computed for the standard regions of interest. The femurs were then categorised according to their BMD-based T-score as osteoporotic, osteopenic and normal. Our dataset counts 28 osteoporotic femurs, 22 osteopenic femurs, and 20 normal femurs. We added this information in the following table:

Table C1
Description of the dataset with age, gender, sides, aBMD values obtained by DXA for the total proximal femur (*Total*), femoral neck (*Neck*), trochanter (*Troch*), and intertrochanter (*Inter*) regions and osteoporotic status based on T-score.

Donor	Age	Gender	Side	BMD				Status
				Total	Neck	Troch	Inter	
1	70	F	L	0.527	0.488	0.411	0.571	Osteoporotic
1	70	F	R	0.449	0.419	0.326	0.521	Osteoporotic
2	60	M	L	0.607	0.594	0.43	0.723	Osteoporotic
2	60	M	R	0.655	0.614	0.498	0.748	Osteoporotic
3	77	F	L	0.645	0.483	0.407	0.821	Osteopenic
3	77	F	R	0.632	0.453	0.385	0.843	Osteopenic
4	84	F	L	0.428	0.31	0.302	0.541	Osteoporotic
4	84	F	R	0.442	0.339	0.272	0.574	Osteoporotic
5	68	F	L	0.892	0.783	0.822	0.972	Normal
5	68	F	R	0.822	0.699	0.677	0.94	Normal

(*continued on next page*)

Table C1 (*continued*)

6	60	M	L	0.69	0.601	0.68	0.712	Osteopenic
6	60	M	R	0.706	0.641	0.648	0.754	Osteopenic
7	80	M	L	0.717	0.625	0.642	0.799	Osteopenic
7	80	M	R	0.738	0.558	0.666	0.829	Osteopenic
8	79	M	L	0.61	0.531	0.495	0.706	Osteopenic
8	79	M	R	0.677	0.55	0.491	0.819	Osteopenic
9	87	F	L	0.444	0.456	0.297	0.521	Osteoporotic
9	87	F	R	0.398	0.397	0.255	0.481	Osteoporotic
10	96	F	L	0.398	0.314	0.245	0.526	Osteoporotic
10	96	F	R	0.282	0.297	0.128	0.39	Osteoporotic
11	66	F	L	0.634	0.556	0.441	0.742	Osteoporotic
11	66	F	R	0.608	0.548	0.437	0.73	Osteoporotic
12	65	M	L					DXA missing
12	65	M	R					DXA missing
13	69	M	L	0.598	0.546	0.401	0.753	Osteoporotic
13	69	M	R	0.544	0.478	0.371	0.671	Osteoporotic
14	89	M	L	1.004	0.811	0.785	1.17	Normal
14	89	M	R	1.01	0.822	0.847	1.176	Normal
15	91	M	L	0.735	0.526	0.53	0.913	Osteopenic
15	91	M	R	0.727	0.533	0.548	0.891	Osteopenic
16	73	F	L	0.806	0.708	0.552	0.992	Osteopenic
16	73	F	R	0.787	0.696	0.555	0.97	Osteopenic
17	73	M	L	0.975	0.806	0.746	1.134	Normal
17	73	M	R	0.894	0.737	0.721	1.041	Normal
18	67	M	L	0.843	0.656	0.667	1.007	Normal
18	67	M	R	0.903	0.673	0.724	1.072	Normal
19	71	M	L	1.18	0.916	0.789	1.434	Normal
19	71	M	R	1.037	0.833	0.707	1.249	Normal
20	88	M	L	0.94	0.791	0.683	1.125	Normal
20	88	M	R	0.949	0.762	0.717	1.126	Normal
21	81	F	L	0.563	0.473	0.376	0.685	Osteoporotic
21	81	F	R	0.588	0.492	0.409	0.699	Osteoporotic
22	85	M	L	0.613	0.497	0.413	0.749	Osteoporotic
22	85	M	R	0.592	0.502	0.37	0.709	Osteoporotic
23	74	F	L	0.628	0.482	0.484	0.793	Osteoporotic
23	74	F	R	0.592	0.432	0.472	0.714	Osteoporotic
24	87	F	L	0.476	0.486	0.334	0.548	Osteoporotic
24	87	F	R	0.467	0.479	0.33	0.547	Osteoporotic
25	64	F	L	0.735	0.656	0.515	0.898	Osteopenic
25	64	F	R	0.767	0.685	0.554	0.934	Osteopenic
26	80	M	L	0.834	0.735	0.638	0.946	Osteopenic
26	80	M	R	0.768	0.698	0.607	0.867	Osteopenic
27	67	M	R	1.009	0.806	0.756	1.201	Normal
27	67	M	L	1.035	0.801	0.771	1.223	Normal
28	93	F	L	0.534	0.437	0.403	0.624	Osteoporotic
28	93	F	R	0.563	0.497	0.401	0.689	Osteoporotic
29	88	F	L	0.399	0.234	0.375	0.448	Osteoporotic
29	88	F	R	0.408	0.284	0.333	0.468	Osteoporotic
30	88	F	L	0.596	0.459	0.495	0.717	Osteoporotic
30	88	F	R	0.586	0.582	0.411	0.674	Osteoporotic
31	59	M	L	0.875	0.714	0.736	0.98	Normal
31	59	M	R	0.889	0.742	0.7	1.032	Normal
32	79	M	L	0.792	0.679	0.51	0.958	Osteopenic
32	79	M	R	0.763	0.703	0.487	0.918	Osteopenic
33	70	F	L	0.831	0.695	0.645	1.018	Normal
33	70	F	R	0.824	0.703	0.631	0.972	Normal
34	89	F	L	0.781	0.543	0.674	0.918	Osteopenic
34	89	F	R	0.733	0.597	0.622	0.84	Osteopenic
35	88	F	L	0.923	0.692	0.661	1.165	Normal
35	88	F	R	0.828	0.676	0.676	0.96	Normal
36	46	F	L	0.797	0.667	0.607	0.905	Osteopenic
36	46	F	R	0.775	0.681	0.6	0.881	Osteopenic

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2018.11.001.

References

Alexander, D.C., Zikic, D., Zhang, J., Zhang, H., Criminisi, A., 2014. Image quality transfer via random forest regression: Applications in diffusion mri. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014: 17th International Conference, Boston, MA, USA, September 14-18, 2014, Proceedings, Part III 225–232. doi:10.1007/978-3-319-10443-0_29.

Arbenz, P., Flaig, C., 2008. On Smoothing Surfaces in Voxel Based Finite Element Analysis of Trabecular Bone. Large-Scale Sci. Comput. (1) 69–77. doi:10.1007/978-3-540-78827-0_6.

Belyaev, A., Ohtake, Y., 2003. A Comparison of Mesh Smoothing Methods. Israel-Korea Bi-National Conference on Geometric Modeling and Computer Graphics 1–5. 10.1.1.61.3683

Blanc, R., Seiler, C., Székely, G., Nolte, L.-P., Reyes, M., 2012. Statistical model based shape prediction from a combination of direct observations and various surrogates: Application to orthopaedic research. Med. Image Anal. 16 (6), 1156–1166. doi:10.1016/j.media.2012.04.004.

Bouatmane, S., Nekhoul, B., Bouridane, A., Tanougast, C., 2007. Classification of prostatic tissues using feature selection methods. In: Jarm, T., Kramar, P., Zupanic, A. (Eds.), 11th Mediterranean Conference on Medical and Biomedical Engineering and Computing 2007. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 843–846.

Bucki, M., Lobos, C., Payan, Y., 2010. A fast and robust patient specific Finite Element mesh registration technique: Application to 60 clinical cases. Med. Image Anal. 14 (3), 303–317. doi:10.1016/j.media.2010.02.003.

Carballido-Gamio, J., Bonaretti, S., Saeed, I., Harnish, R., Recker, R., Burghardt, A.J., Keyak, J.H., Harris, T., Khosla, S., Lang, T.F., 2015. Automatic multi-parametric quantification of the proximal femur with quantitative computed tomography. Quant. Imaging Med. Surgery 5 (4).

Cauley, J.A., 2013. Public health impact of osteoporosis. J. Gerontol. Series A, Biol. Sci. Medical Sci. 68 (10), 1243–1251. doi:10.1093/gerona/glt093.

Cauley, J.A., Cawthon, P.M., Peters, K.E., Cummings, S.R., Ensrud, K.E., Bauer, D.C., Taylor, B.C., Shikany, J.M., Hoffman, A.R., Lane, N.E., Kado, D.M., Stefanick, M.L., Orwoll, E.S., 2016. Risk Factors for Hip Fracture in Older Men: The Osteoporotic Fractures in Men Study (MrOS). J. Bone Mineral Res. 31 (10), 1810–1819. doi:10.1002/jbmr.2836.

Chandran, V., Reyes, M., Zysset, P., 2017. A novel registration-based methodology for prediction of trabecular bone fabric from clinical qct: A comprehensive analysis. PLOS ONE 12 (11), 1–23. doi:10.1371/journal.pone.0187874.

Chandran, V., Zysset, P., Reyes, M., 2015. Prediction of trabecular bone anisotropy from quantitative computed tomography using supervised learning and a novel morphometric feature descriptor. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part I 621–628. doi:10.1007/978-3-319-24553-9_76.

Cootes, T., Taylor, C., Cooper, D., Graham, J., 1995. Active shape models-their training and application. Comput. Vision Image Understand 61 (1), 38–59. doi:10.1006/cviu.1995.1004.

Dall'Ara, E., Luisier, B., Schmidt, R., Kainberger, F., Zysset, P., Pahr, D., 2013. A nonlinear qct-based finite element model validation study for the human femur tested in two configurations in vitro. Bone 52 (1), 27–38. doi:10.1016/j.bone.2012.09.006.

Dall'Ara, E., Eastell, R., Viceconti, M., Pahr, D., Yang, L., 2016. Experimental validation of dxa-based finite element models for prediction of femoral strength. J. Mech. Behav. Biomed. Materials 63, 17–25. doi:10.1016/j.jmbbm.2016.06.004.

Dumont, M., Marée, R., 2009. Fast multi-class image annotation with random windows and multiple output randomized trees. Proc. of VISAPP 2.

Engelke, K., Lang, T., Khosla, S., Qin, L., Zysset, P., Leslie, W.D., Shepherd, J.A., Schousboe, J.T., 2015. Clinical use of quantitative computed tomography (qct) of the hip in the management of osteoporosis in adults: the 2015 {ISCD} official positions—part i. J. Clinical Densitometry 18 (3), 338–358. doi:10.1016/j.jocd.2015.06.012.

Engelke, K., van Rietbergen, B., Zysset, P., 2016. FEA to Measure Bone Strength: A Review. Clinical Rev. Bone Mineral Metabol. 14 (1), 26–37. doi:10.1007/s12018-015-9201-1.

Falcinelli, C., Schileo, E., Pakdel, A., Whyne, C., Cristofolini, L., Taddei, F., 2016. Can ct image deblurring improve finite element predictions at the proximal femur? J. Mech.l Behav. Biomed. Mater. 63, 337–351. doi:10.1016/j.jmbbm.2016.07.004.

Fischler, M.A., Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (6), 381–395.

Geurts, P., Ernst, D., Wehenkel, L., 2006. Extremely randomized trees. Machine Lear 63 (1), 3–42. doi:10.1007/s10994-006-6226-1.

Geuzaine, C., Remacle, J.-F., 2009. Gmsh: A 3-d finite element mesh generator with built-in pre- and post-processing facilities. Int. J. Numer. Methods Eng. 79 (11), 1309–1331. doi:10.1002/nme.2579.

Haralick, R.M., Shanmugam, K., 1973. Textural Features for Image Classification. IEEE TSMC-3 6.

Humbert, L., Hazrati Marangalou, J., del Río Barquero, L.M., van Lenthe, G.H., van Rietbergen, B., 2016. Technical Note: Cortical thickness and density estimation from clinical CT using a prior thickness-density relationship. Med. Phys. 43 (4), 1945–1954. doi:10.1118/1.4944501.

Kang, Y., Engelke, K., Fuchs, C., Kalender, W.a., 2005. An anatomic coordinate system of the femoral neck for highly reproducible BMD measurements using 3D QCT. Comput. Med. Imaging Graph. 29 (7), 533–541.

Kang, Y., Engelke, K., Kalender, W., 2003. A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. IEEE Trans. Med.Imag. 22 (5), 586–598.

Keaveny, T. M., 2017. Virtuost: Bone strength assesment. [Online; accessed 10-September-2017].

Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.W., 2010. elastix: A toolbox for intensity-based medical image registration. IEEE Trans. Med. Imag. 29 (1), 196–205. doi:10.1109/TMI.2009.2035616.

Latypova, A., Maquer, G., Elankumaran, K., Pahr, D., Zysset, P., Pioletti, D.P., Terrier, A., 2016. Identification of elastic properties of human patellae using micro-finite element analysis. J. Biomech. 49 (13), 3111–3115. doi:10.1016/j.jbiomech.2016.07.031.

Lindner, C., Thomson, J., Consortium, T.a., Cootes, T.F., 2015. Learning-based shape model matching: Training accurate models with minimal manual input. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 580–587. doi:10.1007/978-3-319-24574-4_69.

Liu, G.R., Dai, K.Y., Nguyen, T.T., 2007. A smoothed finite element method for mechanics problems. Computat. Mech. 39 (6), 859–877. doi:10.1007/s00466-006-0075-4.

Lorensen, W.E., Cline, H.E., 1987. Marching cubes: A high resolution 3d surface construction algorithm. SIGGRAPH Comput. Graph. 21 (4), 163–169.

Lu, P., Barazzetti, L., Chandran, V., Gavaghan, K., Weber, S., Gerber, N., Reyes, M., 2015. Facial nerve image enhancement from cbct using supervised learning technique. In: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2964–2967. doi:10.1109/EMBC.2015.7319014.

Lu, P., Barazzetti, L., Chandran, V., Gavaghan, K., Weber, S., Gerber, N., Reyes, M., 2017. Highly accurate facial nerve segmentation refinement from cbct/ct imaging using a super resolution classification approach. IEEE Trans. Biomed. Eng..

Luisier, B., Ara, E.D., Pahr, D., 2014. Orthotropic hr-pqct-based {FE} models improve strength predictions for stance but not for side-way fall loading compared to isotropic qct-based {FE} models of human femurs. Journal of the Mechanical Behavior of Biomedical Materials 32, 287–299. doi:10.1016/j.jmbbm.2014.01.006.

Lüthi, M., Gerig, T., Jud, C., Vetter, T., 2017. Gaussian process morphable models. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Lüthi, M., Jud, C., Vetter, T., 2013. A unified approach to shape model fitting and non-rigid registration. Machine Learning in Medical Imaging: 4th International Workshop, MLMI 2013, Held in Conjunction with MICCAI 2013, Nagoya, Japan, September 22, 2013. Proceedings 66–73. doi:10.1007/978-3-319-02267-3_9.

Maquer, G., Buerki, A., Nuss, K., Zysset, P.K., Tannast, M., 2016. Head-neck osteoplasty has minor effect on the strength of an ovine cam-fai model: In vitro and finite element analyses. Clinical Orthopaedics Related Res. 474 (12), 2633–2640. doi:10.1007/s11999-016-5024-8.

Marée, R., Wehenkel, L., Geurts, P., 2013. Extremely randomized trees and random subwindows for image classification, annotation, and retrieval. Decision Forests Comput. Vision Med. Image Anal. 125–141. doi:10.1007/978-1-4471-4929-3_10.

Mayhew, P.M., Thomas, C.D., Clement, J.G., Loveridge, N., Beck, T.J., Bonfield, W., Burgoyne, C.J., Reeve, J., 2005. Relation between age, femoral neck cortical stability, and hip fracture risk. Lancet 366 (9480), 129–135. doi:10.1016/S0140-6736(05)66870-5.

Pahr, D.H., Schwiedrzik, J., Dall'Ara, E., Zysset, P.K., 2014. Clinical versus pre-clinical {FE} models for vertebral body strength predictions. J. Mechan. Behav. Biomed. Mater. 33, 76–83. Forensic Biomechanics, doi: 10.1016/j.jmbbm.2012.11.018.

Pahr, D.H., Zysset, P.K., 2009. From high-resolution ct data to finite element models: development of an integrated modular framework. Comput. Methods Biomech. Biomed. Eng. 12 (1), 45–57. doi:10.1080/10255840802144105. PMID: 18839383

Pakdel, A., Robert, N., Fialkov, J., Maloul, A., Whyne, C., 2012. Generalized method for computation of true thickness and x-ray intensity information in highly blurred sub-millimeter bone features in clinical CT images. Phys. Med. Biol. 57 (23), 8099–8116. doi:10.1088/0031-9155/57/23/8099.

Pauchard, Y., Fitze, T., Browarnik, D., Eskandari, A., Pauchard, I., Enns-Bray, W., Pálsson, H., Sigurdsson, S., Ferguson, S.J., Harris, T.B., Gudnason, V., Helgason, B., 2016. Interactive graph-cut segmentation for fast creation of finite element models from clinical ct data for hip fracture prediction. Comput. Methods Biomechan. Biomed, Eng. 19 (16), 1693–1703. PMID: 27161828, doi: 10.1080/10255842.2016.1181173.

Peleg, E., Herblum, R., Beek, M., Joskowicz, L., Liebergall, M., Mosheiff, R., Whyne, C., 2014. Can a partial volume edge effect reduction algorithm improve the repeatability of subject-specific finite element models of femurs obtained from ct data? Comput. Methods Biomech. Biomed. Eng. 17 (3), 204–209. PMID: 22452517, doi: 10.1080/10255842.2012.673595.

Poole, K.E., Skingle, L., Gee, A.H., Turmezei, T.D., Johannesdottir, F., Blesic, K., Rose, C., Vindlacheruvu, M., Donell, S., Vaculik, J., Dungl, P., Horak, M., Stepan, J.J., Reeve, J., Treece, G.M., 2017. Focal osteoporosis defects play a key role in hip fracture. Bone 94, 124–134. doi:10.1016/j.bone.2016.10.020.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P., 2007. Numerical Recipes 3rd Edition: The Art of Scientific Computing, 3 Cambridge University Press, New York, NY, USA.

Prevrhal, S., Engelke, K., Kalender, W.a., 1999. Accuracy limits for the determination of cortical width and density: the influence of object size and CT imaging parameters. Phys. Med.Biol. 44 (3), 751–764. doi:10.1088/0031-9155/44/3/017.

Prevrhal, S., Fox, J.C., Shepherd, J.A., Genant, H.K., 2003. Accuracy of CT-based thickness measurement of thin structures: modeling of limited spatial resolution in all three dimensions. Med. Phys. 30 (1), 1–8. doi:10.1118/1.1521940.

Scalismo, 2016. scalismo - scalable image analysis and shape modelling. https://www.github.com/unibas-gravis/scalismo.

Shahim, K., Jürgens, P., Cattin, P.C., Nolte, L.-P., Reyes, M., 2013. Prediction of cranio–maxillofacial surgical planning using an inverse soft tissue modelling approach. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 18–25.

Stone, K.L., Seeley, D.G., Lui, L.-Y., Cauley, J.A., Ensrud, K., Browner, W.S., Nevitt, M.C., Cummings, S.R., 2003. Bmd at multiple sites and risk of fracture of multiple types: Long-term results from the study of osteoporotic fractures. J. Bone Mineral Res. 18 (11), 1947–1954. doi:10.1359/jbmr.2003.18.11.1947.

Taghizadeh, E., Chandran, V., Reyes, M., Zysset, P., Büchler, P., 2017. Statistical analysis of the inter-individual variations of the bone shape, volume fraction and fabric and their correlations in the proximal femur. Bone.

Taubin, G., 1995. Curve and surface smoothing without shrinkage. Proc. IEEE Int. Conf. Comput. Vision 852–857. doi:10.1109/ICCV.1995.466848.

Treece, G., Gee, A., 2015. Independent measurement of femoral cortical thickness and cortical bone density using clinical {CT}. MedIA 20, 249–264.

Treece, G., Poole, K., Gee, A., 2012. Imaging the femoral cortex: Thickness, density and mass from clinical {CT}. Med. Image Anal. 16 (5), 952–965. doi:10.1016/j.media.2012.02.008.

Treece, G.M., Gee, a.H., Mayhew, P.M., Poole, K.E.S., 2010. High resolution cortical bone thickness measurement from clinical CT data. Med. Image Anal. 14 (3), 276–290. doi:10.1016/j.media.2010.01.003.

Treece, G.M., Gee, A.H., Tonkin, C., Ewing, S.K., Cawthon, P.M., Black, D.M., Poole, K.E., for the Osteoporotic Fractures in Men (MrOS) Study, ., 2015. Predicting hip fracture type with cortical bone mapping (cbm) in the osteoporotic fractures in men (mros) study. J. Bone Mineral Res. 30 (11), 2067–2077. doi:10.1002/jbmr.2552.

Varga, P., Baumbach, S., Pahr, D., Zysset, P., 2009. Validation of an anatomy specific finite element model of colles' fracture. J. Biomech. 42 (11), 1726–1731. doi:10.1016/j.jbiomech.2009.04.017.

Whitmarsh, T., Treece, G.M., Gee, A.H., Poole, K.E.S., 2016. The effects on the femoral cortex of a 24 month treatment compared to an 18 month treatment with teriparatide: A multi-trial retrospective analysis. PLOS ONE 11 (2), 1–9. doi:10.1371/journal.pone.0147722.

Zhang, J., Hislop-Jambrich, J., Besier, T.F., 2016. Predictive statistical models of baseline variations in 3-d femoral cortex morphology. Med. Eng. Phys. 38 (5), 450–457. doi:10.1016/j.medengphy.2016.02.003.

Zhou, S.K., Comaniciu, D., 2007. Shape regression machine. In: Karssemeijer, N., Lelieveldt, B. (Eds.), Information Processing in Medical Imaging. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 13–25.

Zysset, P., Pahr, D., Engelke, K., Genant, H.K., McClung, M.R., Kendler, D.L., Recknor, C., Kinzl, M., Schwiedrzik, J., Museyko, O., Wang, A., Libanati, C., 2015. Comparison of proximal femur and vertebral body strength improvements in the {FREEDOM} trial using an alternative finite element methodology. Bone 81, 122–130. Epigenetic Mechanisms Regulating Bone Biology and Pathology, doi: 10.1016/j.bone.2015.06.025.