IEEE Access
The journal for rapid open access publishing

Received August 4, 2015, accepted September 11, 2015, date of publication September 30, 2015,
date of current version October 23, 2015.

Digital Object Identifier 10.1109/ACCESS.2015.2484259

A New Level-Set-Based Protocol for Accurate Bone Segmentation From CT Imaging

MANUEL PINHEIRO AND J. L. ALVES
Department of Mechanical Engineering, University of Minho, Guimarães 4800-058, Portugal
Corresponding author: M. Pinheiro (manuelspinheiro@gmail.com)

This work was supported by the Fundação para a Ciência e Tecnologia (FCT), Portugal, under Project UID/EEA/04436/2013. The work of M. Pinheiro was supported by FCT for the Ph.D. under Grant SFRH/BDE/51143/2010.

ABSTRACT A new medical image segmentation pipeline for accurate bone segmentation from computed tomography (CT) imaging is proposed in this paper. It is a two-step methodology, with a pre-segmentation step and a segmentation refinement step, as follows. First, the user performs a rough segmenting of the desired region of interest. Second, a fully automatic refinement step is applied to the pre-segmented data. The automatic segmentation refinement is composed of several sub-steps, namely, image deconvolution, image cropping, and interpolation. The user-defined pre-segmentation is then refined over the deconvolved, cropped, and up-sampled version of the image. The performance of the proposed algorithm is exemplified with the segmentation of CT images of a composite femur bone, reconstructed with different reconstruction protocols. Segmentation outcomes are validated against a gold standard model, obtained using the coordinate measuring machine Nikon Metris LK V20 with a digital line scanner LC60-D and a resolution of 28 μm. High sub-pixel accuracy models are obtained for all tested data sets, with a maximum average deviation of 0.178 mm from the gold standard. The algorithm is able to produce high quality segmentation of the composite femur regardless of the surface meshing strategy used.

INDEX TERMS Biomedical image processing, deconvolution, image segmentation, level set, spatial resolution.

I. INTRODUCTION

The first milestone towards custom implant development is the accurate bone segmentation from medical image data. The evolution of Computed Tomography (CT) imaging led to the widespread of this technique to all medical fields. From the engineering standpoint, CT imaging can be used in the development of patient-specific biomechanical and finite element models, as well as in the development of custom implants [1]–[4].

Currently, CT imaging is the modality of choice to image the Human skeletal system. This technique has the ability to enhance the radiological contrast between hard and soft-tissue, which simplifies image segmentation, and allows the production of accurate representations of the bone. The accurate segmentation of bone is important not only to guarantee the overall fit to the patient's anatomy, but also for the success of the implant [5], [6]. However, the degree of anatomical fit necessary to minimize the biological impact of the implant is still unknown. On the one hand, too much implant fit may preclude implant insertion, and may damage to the host bone [7]. On the other hand, the absence of implant fit may cause interfacial micromotions, which prevent bone ingrowth and implant osseointegration [8], [9].

During image acquisition, the CT scanner acts as a low-pass filter. Due to the limited frequency response of the imaging system, an overall degradation on image quality is observed. The sharp transitions between adjacent structures become diffuse in the final image. Near boundaries the attenuation coefficients of two (or more) adjacent structures are averaged, a phenomenon often called partial volume effect. Partial volume effect may be detrimental to the accuracy of the segmentation process. In [10] it was concluded that CT imaging produced large domain overestimations, for structures with a cortical thickness smaller than 2.0 mm.

The amount of spatial blurring in an image is often modelled by the system's Point Spread Function (PSF). The PSF is also thought to be a practical measure of the spatial resolution of the CT scanner. In [11] it was concluded that the inner and the outer cortical surfaces of bone could only be accurately determined for thickness greater than the Full Width at Half Maximum (FWHM) of the PSF. In addition,

for objects with a thickness smaller than the FWHM of the PSF, domain overestimations up to 40% of their original size were obtained in [12]. Similarly, in [13] it was observed that for small diameter spheres imaged with CT the apparent diameter was in fact the FHWM of the PSF. Therefore, the FWHM of the PSF seems to provide a measure of the maximum spatial frequency that can be accurately encoded by a given CT machine, rather than other acquisition settings.

The Field of View (FOV) used during image reconstruction may also affect the overall accuracy of the acquisition process. Small FOV are normally used in high-resolution reconstructions, and the PSF if often the limiting factor to the spatial resolution. However, when imaging large FOV, the size of the reconstructed pixel may be larger than the FWHM of the PSF, and the PSF spreads to occupy a single pixel in the reconstructed image [14]. Image reconstruction with large voxel sizes is highly detrimental for the accurate representation of the bone's surface, especially for thinner and highly curved structures [15].

The image acquisition process may also be affected by several image artifacts, such as noise, beam hardening effect, motion artifacts, among others [16]. These shortcomings, associated with the image acquisition process, may be minimal with respect to other error sources. In fact, in CT imaging the image reconstruction is known to have a very high accuracy, and to be almost free of geometrical magnification [17]. Contrarily to the acquisition and reconstruction processes, image segmentation is often affected by high inter and intra-expert variability, and the image processing and segmentation chain may contribute with up to 70% of the average error found in the final reconstruction [18].

In literature, there are numerous studies addressing bone segmentation, but only a few of them evaluate the accuracy of the segmentation outcome. In an early study, Rothuizen et al. [19] analysed the CT attenuation profile normal to the bone's surface, and concluded that a single threshold was insufficient to accurately define the femur's cortical shell. Two relative thresholds were proposed to segment the diaphyseal and metaphyseal regions of the femur properly (45% and 50% of the maximum HU value along each individual attenuation profile, respectively). More recently, Kang et al. [12], Aamodt [20], and Rathnayaka et al. [21] reported the achievement of accurate segmentations of cortical bone structures from CT imaging with single thresholding, adaptive thresholding, and multiple thresholding, respectively. In [22] a Levenberg-Marquardt-based relative thresholding segmentation algorithm was proposed. The algorithm fits an ideal high-resolution attenuation model to the HU attenuation profile, normal to the bone's surface, to obtain sub-millimetre estimates of the cortical thickness. Bone segmentation is obtained applying 50% relative threshold to the fitted ideal HU attenuation profile. A similar approach was employed in [23], to segment thin craniofacial bones. In both [22] and [23], sub-pixel accuracy estimates of the inner and outer surface of bone were reported.

In practice, the aforementioned techniques may display some limitations. Single thresholding is very sensitive to image inhomogenities, noise, and threshold selection. In addition, adaptive thresholding does not guarantee the production of closed contours, and additional manual segmentation may be needed to obtain topologically correct segmentations [12]. Adaptive thresholding may also provide incorrect estimates of the bone surface, especially in the presence of thin cortical shells [11]. The application of adaptive/relative thresholding together with model fitting techniques seems to overcome the aforementioned limitations. However, the application of these techniques to each boundary pixel is extremely time consuming and unpractical. In addition, the normal direction is highly affected by the discrete nature of the image [24]. The surface of the bone has to be sampled and each surface point needs to be processed independently, which produces highly irregular contours. It also produces unreliable estimates of the bone surface if the observed attenuation profile deviates from the ideal attenuation profile, which is common near the articulating surfaces of bones.

In this work a different approach to an accurate bone segmentation is proposed. The proposed protocol allows the segmentation of bony structures with sub-pixel accuracy, and intrinsically guarantees the smoothness of the extracted contours. The proposed segmentation protocol is validated through the comparison between the segmentation outcome and a geometrically well-defined gold standard. For validation purposes, a synthetic bone was used as gold standard. The remainder of this paper is organized as follows: in section II the segmentation protocol, the definition of the segmentation gold standard, and the means for quantifying the accuracy of the CT machine are described; in section III one presents a methodology to estimate the PSF of the CT machine; in section IV the impact of the domain discretization (voxel vs. average error) in the model accuracy is shortly analysed; section V and section VI refer to the description and discussion of the results obtained with the proposed segmentation pipeline; and in section VII are drawn and discussed the main conclusions of the present work.

II. MATERIALS AND METHODS

General purpose image segmentation protocols are often affected by the variability in the Human anatomy, the image acquisition process, image artifacts, among others. Commonly, segmentation methods designed specifically to each application often produce better results than general purpose algorithms. Nevertheless, some degree of standardization is desirable, particularly when the segmentation is part of the product development pipeline. Therefore, in this work it is proposed a new two-step segmentation protocol for reliably and accurately extract hard tissue structures from image data. To evaluate the accuracy of the newly proposed segmentation protocol, a phantom study is carried out. One composite replica of the Human femur (Fig. 1 (a)) commercially

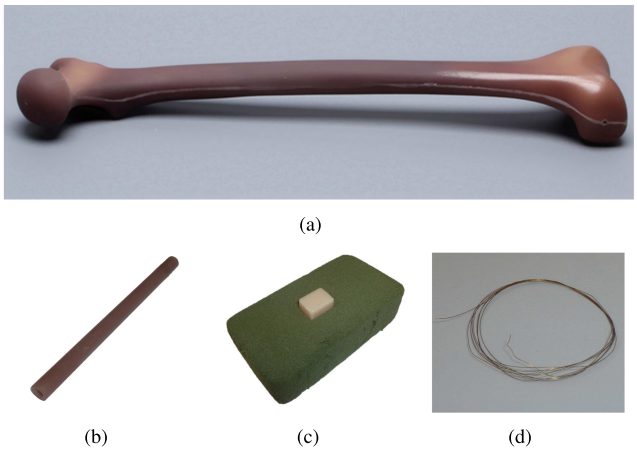


FIGURE 1. In (a) the gold standard composite femur and in (b), (c) and (d) the reference calibrated hollow cylinder with 9.81 mm of diameter, the ceramic box and the brass alloy wire with 0.10 mm of diameter used to estimate the system's Point Spread Function.

available at the Sawbones website[1] was imaged with a CT machine. The image acquisition process was carried with a fourth-generation CT scanner Toshiba Aquilion^{TM} 64 at the CUF Hospital, Porto (Portugal). Several image datasets from the composite femur were obtained, with two distinct acquisition protocols. The accuracy of the segmentation was validated against a gold standard model, obtained using the coordinate measuring machine Nikon Metris LK V20 with a digital line scanner LC60-D that guarantees an accuracy of 28 μm.

For evaluation purposes, the outcome of each segmentation is compared with the gold-standard, using two distance measures, namely the Mean Symmetric Distance (MSD) and the Hausdorff Distance (HD) [25]. The Mean Symmetric Distance (MSD) quantifies the average distance between each point of the segmentation model S and the closest point of the reference model R (1):

$$MSD(S,R) = \frac{1}{m+n}\left[\sum_{i=1}^{m}\min_{S_i}\left(d\left(S_i,R_j\right)\right) + \sum_{j=1}^{n}\min_{R_j}\left(d\left(R_j,S_i\right)\right)\right] \quad (1)$$

[1]http://www.sawbones.com/

where m and n stand for the number of surface points in the segmentation S, and the number of surface points in the gold-standard model R, respectively. The Hausdorff Distance (HD) computes the maximum deviation between the two models (2):

$$HD(S,R) = \max\left[\min_{S_i}\left(d\left(S_i,R_j\right)\right), \min_{R_j}\left(d\left(R_j,S_i\right)\right)\right] \quad (2)$$

where again R stands for the gold-standard model, and S stands for the segmented model.

Three geometrically simpler phantom objects were added to the scanning process, to estimate the limiting resolution of the CT machine. In practice, these geometrically simpler phantoms may be important to characterize the image acquisition process, especially the limiting resolution of the CT scanner. The objects tested were a calibrated hollow cylinder with an outer diameter of 9.81 ± 0.02 mm, a ceramic box with dimensions $12.51 \times 13.81 \times 18.01 \pm 0.01$ mm, and a brass alloy wire with a diameter of 0.10 mm (Fig. 1 (b), (c) and (d)). The compact nature of these phantom objects allow them to be imaged simultaneously with the patient, and may avoiding the need to have a dedicated phantom to assess the spatial resolution of the CT scanner. The ability to quantify the system's limiting resolution was evaluate against the CATPHAN 528, which is a commercially available phantom often used for quality control [26]. The CATPHAN 528 was imaged to quantify the true in-plane resolution of the CT machine, against which the performance of each phantom could be compared.

Regarding CT scanning, different machine settings were used to the acquisition and reconstruction processes. The summary of image acquisition protocols and image reconstruction resolutions for the phantom femur, the CATPHAN 528, and the three other phantoms is presented in Table 1. The images in Dataset (DS) #1 were obtained with the highest possible spatial resolution of the scanner. In Dataset #2 and #3 the standard settings for imaging the hip and femur were used in the acquisition process. Two additional Datasets were computed from the high resolution DS #1, in order to test the segmentation protocol in low resolution scans (DS #4 and DS #5). The CATPHAN 528

TABLE 1. CT image acquisition protocol summary and target reconstruction resolution for each Dataset (DS): Datasets #1, #2 and #3 were obtained from the raw data, whereas in Datasets #4 and #5 the in-plane resolution was downscaled to 1/2 and 1/4 of the original (reconstructed) resolution.

	Data Acquisition								
	DS#1	DS#2	DS#3	Cylinder	Box	CATPHAN 528	Wire	DS#4	DS#5
Tube Voltage (kV)	120	120	120	120	120	120	120	120	120
Tube Current (mA)	200	200	200	200	200	200	200	200	200
Scan FOV (mm)	240	400	400	240	400	400	400	240	240
Slice Thickness (mm)	0.3	0.3	3.0	0.3	0.3	0.3	0.3	0.5	1.0
Reconstruction Pixel Size (mm)	0.243	0.525	0.525	0.243	0.525	0.460	0.460	0.486	0.972
Image Matrix	512x512	512x512	512x512			512x512		256x256	128x128
Number of Slices	1547	1671	168					929	465

IEEE Access
The journal for rapid open access publishing

phantom was imaged with the settings used for routine quality control of the CT scan.

To segment the abovementioned Datasets, a two-step segmentation pipeline is proposed. The main reasons for this two-step approach are: i) handling the variability of the Human anatomy in a straightforward manner, and the fact that image segmentations designed for a specific application provide better results than general purpose ones; ii) handling the variability of the acquisition process, which can add different artifacts to the image; iii) handling the variability of the segmentation process across individuals, and guarantee accuracy across different users; and iv) provide some degree of standardization to introduce the image segmentation protocol in an implant development pipeline. Therefore, the segmentation process was divided into two-steps: first, a pre-segmentation step; second, a segmentation refinement step.

In the first step, a pre-segmentation of the image $I(x, y)$ must be provided. In this step the technique(s) more suitable to provide an initial segmentation of the bone may be applied. The output of the pre-segmentation should provide a set of Regions of Interest (ROI) (e.g., the segmentation of the composite femur and the phantom objects), and may also have some spatial constraints (such as the CT table). The spatial constraints aim to identify image pixels that belong to the background, and avoid their segmentation during the refinement step. So, the pre-segmentation step provides some high-level information about the desired domain, as well as some spatial relationships between any existing adjacent structures. One has found that this allows handling segmentation variability more easily during the segmentation refinement process. In the second step, a fully automatic segmentation refinement composed by several sub-steps is performed in order to optimize the initial partition. Fig. 2 schematically depicts the proposed segmentation protocol. The refinement protocol comprises image deconvolution, which is applied in order to minimize the partial volume effect caused by the PSF during image acquisition. Next, the image is cropped around the ROI according to the pre-segmented data, and up-sampled with cubic spline interpolation (3):

$$\beta^3 = \begin{cases} \frac{2}{3} - \frac{1}{2}|x|^2(2-|x|), & 0 \le |x| < 1 \\ \frac{1}{6}(2-|x|)^3, & 1 \le |x| < 2 \\ 0, & 2 \le |x| \end{cases} \quad (3)$$

where x defines the finite support of the basis function. In [27] concluded that cubic spline interpolation provides the best interpolation strategy for image processing applications, both in terms of computation effort and image induced artifacts. Segmentation refinement is performed over the interpolated ROI. The final segmentation obtained using the Active Contours Without Edges (ACWE), proposed by Chan and Vese in [28], and considering the spatial constraints

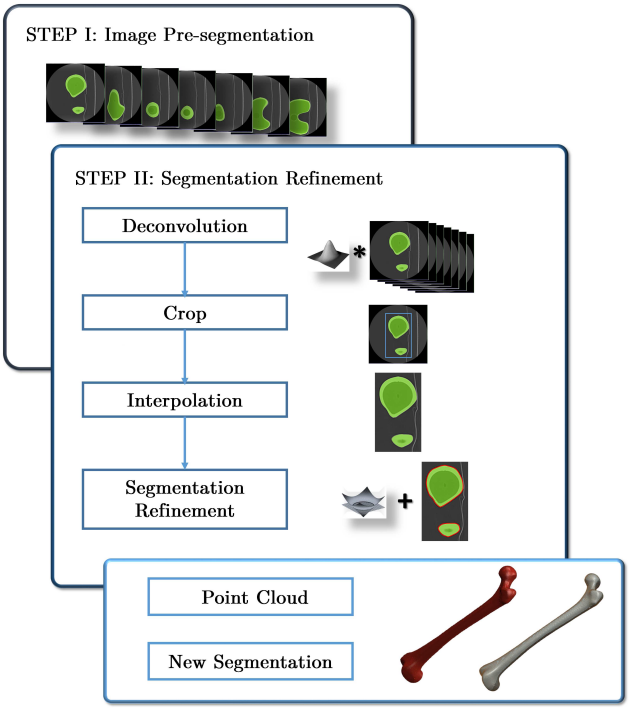


FIGURE 2. Schematic description of the image segmentation protocol proposed: in a first step the user performs a pre-segmentation of the domain that aims to provide some high level information of the desired ROI and adjacent structures; in a second step the pre-segmentation provides the starting point for a fully automatic segmentation refinement which encompasses image deconvolution, image cropping, interpolation and segmentation by the level set method.

defined in the pre-segmentation step (4):

$$\frac{\partial \phi_i}{\partial t} = \tau div\left(\frac{\nabla \phi_i}{|\nabla \phi_i|}\right) - \lambda_1 (I-\mu_1)^2 + \lambda_2 (I-\mu_2)^2 + \eta\mu_2 \sum_{k=1;k\neq i}^{n} S_k \quad (4)$$

where ϕ_i is the level-set corresponding to the pre-segmented region S_i, $\lambda_1 = \lambda_2 = 1$, μ_1, and μ_1 and μ_2 are the average intensity inside and outside the curve. The right hand summation assigns the average intensity of the background to all the remaining pre-segmented sub-regions S_k and spatial constrains defined in the pre-segmentation step, multiplied by a user defined cost η. The ACWE are less sensitive to curve initialization and noise than other level-set methods. The pre-segmentation step is more flexible, i.e. the pre-segmentation may be a rough approximation to the target domain.

The segmentation pipeline was tested in relatively large datasets (see Table 1). In addition, image interpolation increases even further the computational complexity of the segmentation process, especially in terms of memory requirements. To reduce the computation requirements of the algorithms, several simplifications were considered: i) the algorithm is implemented in two dimensions (2D); ii) the segmentation refinement is performed independently for each region, instead of applying it to all regions at the same time (for instance, see the multi-phase level-set method in [29]);

and iii) curve evolution is implemented using the sparse field method [30]. Simplifications i) and ii) help to reduce drastically the computational complexity and memory requirements of the segmentation refinement, by allowing the application of image interpolation and curve evolution in smaller ROI at each iteration.

Two refinement outputs were considered, as schematically depicted in Fig. 2. The new segmentation can be directly exported to the CAD modelling software via point cloud, or downscaled to the original image resolution for further surface or Finite Element mesh generation. To test the robustness against image noise, the segmentation refinement was tested over Datasets #1 to #5 corrupted with Additive White Gaussian Noise (AWGN). In next section the first image deconvolution is addressed. In section III, the performance of phantom object (Fig. 1 (b), (c) and (d)) is evaluated against the CATPHAN 528.

III. IMAGE DECONVOLUTION

It is known from image acquisition theory that any image obtained from an imaging system is not perfect and is only an approximation to the real (ideal) image. The real image is lost due to the intrinsic nature of the acquisition process. However, an estimate of its true distribution may be obtained, considering the output image and some prior knowledge about the system's behaviour. In 2D image acquisition theory, the imaging system is commonly considered as being linear and spatially invariant, and the output image $G(x, y)$ may be correlated with the input image $I(x, y)$ according to (5):

$$G(x, y) = I(x, y) \otimes h(x, y) + n(x, y) \tag{5}$$

where $\otimes$ denotes the 2D convolution, $h(x, y)$ denotes the system blurring effect of the system's PSF, and $n(x, y)$ is an addictive noise term [31]. The Deconvolution problem is intrinsically limited by the knowledge about the PSF and the noisy processes related with the acquisition itself. The noise term $n(x, y)$ is typically a stochastic process that may be originated by a multitude of processes [32]. Noise can be efficiently suppressed by non-linear filtering, for instance applying anisotropic diffusion as proposed in [33]. It also avoids edge bias, which is very common with linear filtering [34], [35].

The PSF is defined by the overall behaviour of the image acquisition system. In practice, the PSF of a given imaging system is frequently approximated by a normalized Gaussian function as stated by the central slice theorem, such that (6):

$$h(x, y, z) = \frac{1}{(2\pi)^{\frac{3}{2}} \sigma_x \sigma_y \sigma_z} e^{\left(\frac{x^2}{2\sigma_x} + \frac{y^2}{2\sigma_y} + \frac{z^2}{2\sigma_z}\right)} \tag{6}$$

where σ_x, σ_y, and σ_z denote the standard deviation in each orthogonal direction. Two assumptions are commonly found in literature for the PSF, namely: (i) the PSF is assumed to be uniformly invariant in the slice plane; and (ii) the cross-plane PSF is generally also assumed to be invariant in the axial direction [11], [22], [36]. In practice, the PSF is not completely isotropic and shift invariant, however such approximation can be safely made for most CT scanners, as well as for several other medical acquisition modalities [37]. For simplicity, one will only consider the estimation of the in-plane blur, and hence (6) can be reduced to its 2D counterpart.

There are several approaches to determine the PSF of an imaging system proposed in the literature. The PSF can be determined: i) observing the impulse response of the system; and ii) computing the system's response to strong edges or Edge Spread Function (ESF). The ESF can be computed simply by differentiating the edge response to radiopaque objects [38], [39].

The phantom objects investigated here allow the measurement of PSF directly (Fig. 1 (d)), or by computing the ESF (Fig. 1 (b) and (c)). The ESF can be easily determined fitting a Gaussian function to the attenuation profile normal to the objects surface. The ESF was determined according to the following procedure: i) the reference CAD model was superimposed to the image data; ii) computation of the gradient near object boundaries; iii) 1D profile de-noising using Wavelet thresholding as proposed in [40]. Recently, the *SureShrink* Wavelet thresholding proved to be extremely efficient in removing noise, when compared with more recent techniques [41]; and iv) curve fitting over a narrow band around the gradient maximum, in order to remove the influence of the adjacent structures. A similar approach was applied in [42] to completely characterize the PSF using a single image. The final estimate of the PSF was defined as the average of all PSF estimates along the sampled phantom edges.

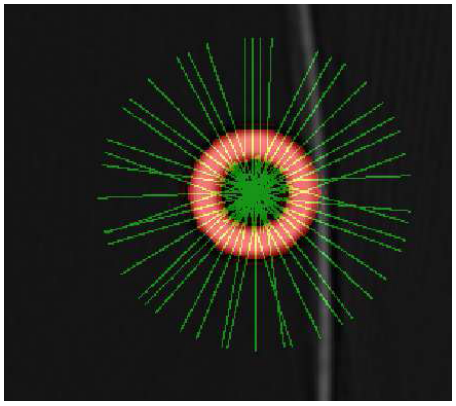
Fig. 3 depicts the PSF of the Toshiba Aquilion^{TM} 64 CT scanner obtained with the CATPHAN 528, and the results obtained with the wire phantom, the hollow cylinder with 9.81 mm of diameter and the ceramic box.

In case of the CATPHAN 528, the PSF has a standard deviation of $\sigma_{528} = 0.3633$ mm and a FWHM of approximately 0.86 mm; with the wire phantom a very similar estimate was obtained, $\sigma_{wire} = 0.3737$ mm and FWHM of 0.88 mm; for the ceramic phantom, the standard deviation of the scanner PSF is slightly overestimated ($\sigma_{box=0.4272}$ mm), and underestimation in case of the hollow cylinder ($\sigma_{cylinder} = 0.3145$ mm). The ceramic box is the proposed phantom that most deviates from the standard deviation of the CATPHAN 528.

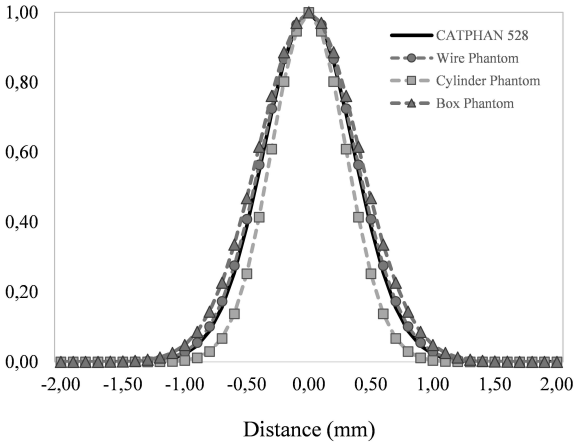
Different deconvolution algorithms were tested, although the best results were obtained using a standard iterative blind deconvolution algorithm, namely the Lucy-Richardson. The Lucy-Richardson algorithm to obtain the new estimates of the original scene $\hat{I}_{k+1}(x, y)$ and the PSF $\hat{h}_{k+1}(x, y)$ are defined as (7):

$$\hat{I}_{k+1}(x, y) = \hat{I}_k(x, y) \left[\frac{G(x, y)}{\hat{h}_k(x, y) \otimes \hat{I}_{k+1}(x, y)} \otimes \hat{h}_k^*(x, y) \right] \tag{7a}$$

$$\hat{h}_{k+1}(x, y) = \hat{h}_k(x, y) \left[\frac{G(x, y)}{\hat{h}_k(x, y) \otimes \hat{I}_{k+1}(x, y)} \otimes \hat{I}_k^*(x, y) \right] \tag{7b}$$

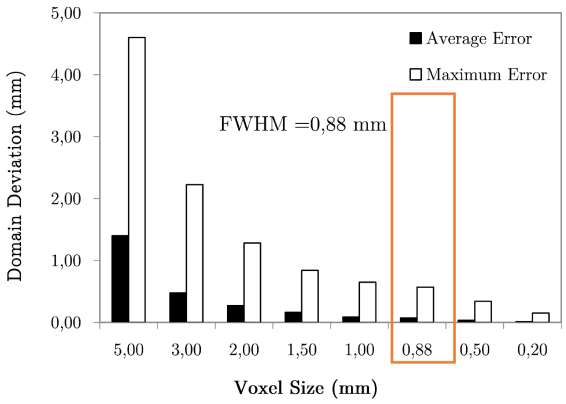


(a)

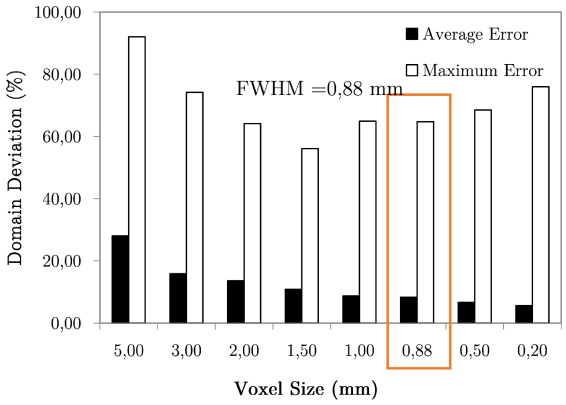


(b)

FIGURE 3. (a) Attenuation profiles (green) normal to the surface of the hollow cylinder (red) to estimate the PSF; (b) the PSF of the fourth-generation Toshiba Aquilion^{TM} 64 CT scanner with the CATPHAN 528, the 0.1 mm wire phantom, and the calibrated hollow cylinder and ceramic box phantoms.



(a)



(b)

FIGURE 4. (a) average (MSD) and maximum (HD) surface errors between the Nikon Metris LK V20 gold standard and the surface mesh obtained after domain discretization with the tested voxel sizes and with a voxel size consistent with the FWHM of the PSF as determined in section III; (b) average and maximum deviation normalized by the voxel size for all voxel sizes.

where $\hat{h}_k^*(x, y)$ and $\hat{I}_k^*(x, y)$ are the complex conjugates of $\hat{h}_k(x, y)$ and $\hat{I}_k(x, y)$, respectively, $k \in N_0$ and $\hat{I}_0(x, y) = G(x, y)$ is the acquired image with image noise $n(x, y)$ being neglected. The blind deconvolution is not dependent on the knowledge of the system's spatial blurring [32], [43]. However, we found that robust results may be obtained if an accurate initial guess of the system's PSF is provided. An implementation of the Lucy-Richardson can be found in MATLAB routine *deconvblind*. This function was used throughout this work for image restoration.

IV. DISCRETIZATION IN DOMAIN ACCURACY

The system's PSF may not be the only factor affecting the resolution of the CT scan. Image reconstruction with large slice thicknesses is commonly associated with larger model inaccuracies [11], [43], [44]. In [45] it was demonstrated that alongside with the slice thickness, the FOV also influences the spatial resolution of the scan and the amount of partial volume averaging. It was concluded that the size of the voxel may be more critical than other error sources, such as the surface meshing parameters. In addition, in [46] the impact of the voxel size and meshing parameters were analysed in the representation of a Human lumbar motion segment. It was also concluded that the resolution of the CT scan (voxel size and slice thickness) was the major source of geometrical inaccuracies of the reconstructed model. Therefore, the effect of domain discretization (slice thickness and voxel size) cannot be neglected, when assessing the achievable model accuracy from a given set of CT images.

To understand the effect of pixel size in domain accuracy, the gold-standard composite femur was discretized with different isotropic voxels sizes, namely {5.0, 3.0, 2.0, 1.5, 1.0, 0.5, 0.2} (mm).

Fig. 4 shows the average and maximum domain error as a function of the voxel size. The average error due to the domain discretization varied between 1.401 ± 0.836 mm and 0.011 ± 0.012 mm, whereas the maximum deviation was 4.603 mm and 0.152 mm for a voxel size of 5.0 mm and 0.2 mm, respectively. The results show that as

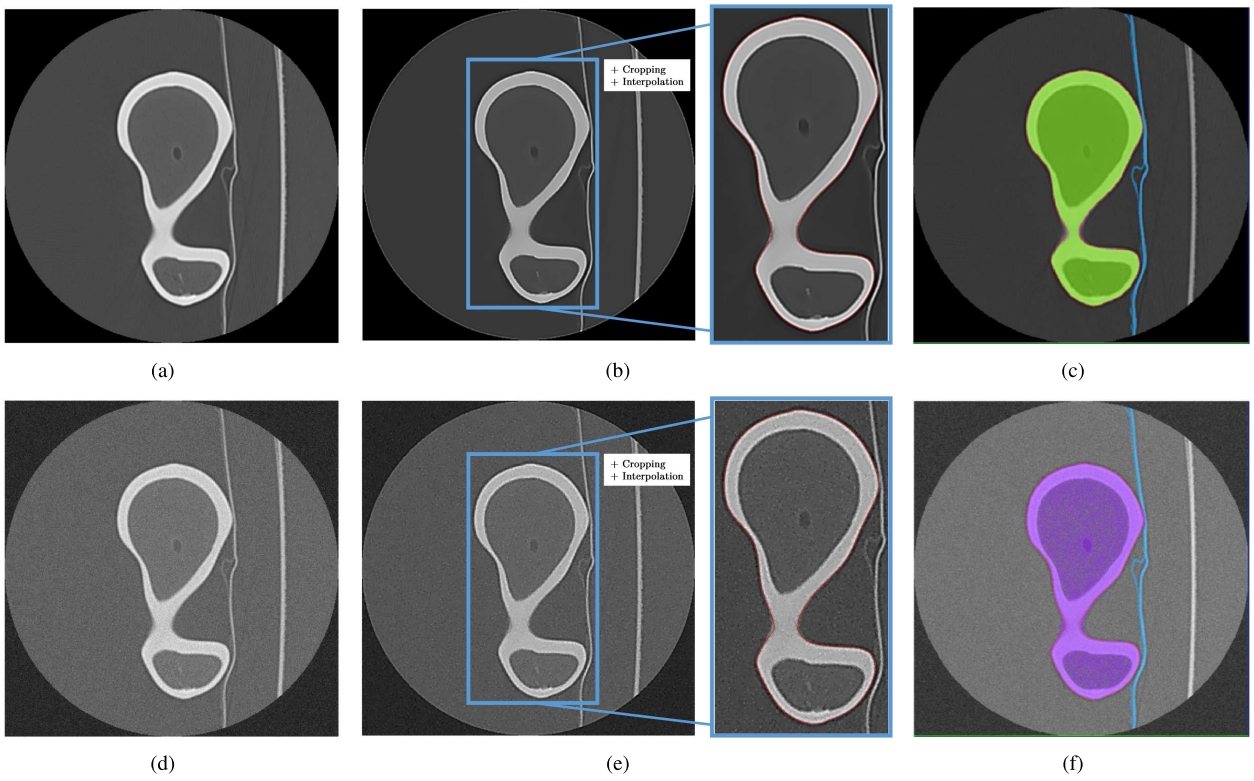


FIGURE 5. Segmentation refinement pipeline applied to the Dataset #1: (a) the original image data without noise; (b) the image data after de-noising with anisotropic diffusion, image deconvolution, and image cropping and interpolation around the pre-segmented region; and (c) the final segmentation contour superimposed over the pre-segmentation mask; (d), (e) and (f) the image refinement is applied to the same image corrupted with AWGN.

the domain sampling increases both the absolute average and the maximum error decrease (Fig. 4 (a)). The largest domain deviations are found in small localized sharp features or surface irregularities along the reference model, which are lost or smoothed due to the domain sampling. For a voxel size consistent with the FWHM of the PSF (0.88 mm) previously determined in section III (with the wire phantom), an average error of 0.073 ± 0.066 mm and a maximum deviation of 0.569 mm are expected to occur, due to the domain discretization.

In Fig. 4 (b) the average and maximum errors are normalized against the sampling size. The normalized average error decreases from 28% to 5.6% of the sampling size, while the maximum error decreases from 92.1% to 56.1% for domain samplings between 5.0 mm and 1.5 mm. However, for edge lengths equal and below 1.0 mm the maximum error shows a growth tendency, reaching 76.0% of the domain sampling for 0.2 mm. The normalized average error decreases and becomes almost constant for smaller sampling sizes. The maximum normalized error increases initially for larger sampling sizes, and increases as the sampling size becomes smaller. Interestingly, for smaller voxel sizes the average error converges to a fixed percentage of the sampling size (5% of the voxel edge size).

For a domain sampling of 0.88 mm (FWHM of the PSF determined in section III) the average error due to discretization is approximately 8% of the sampling size, therefore image interpolation of the image beyond this value may not add any additional information to the segmentation.

V. IMAGE SEGMENTATION AND REFINEMENT RESULTS

Two outputs may be produced after segmentation refinement: i) a high-resolution point cloud model obtained directly from the interpolated image data, and ii) a down-scaled version of the high-resolution segmentation (down to the original resolution of the CT scan). The surface mesh model obtained from the point cloud will be hereafter referred as Point Cloud (PC) model. Since image refinement is performed independently in each image, the PC model is produced directly from a set of equally spaced contours by tiling the cloud points. The down-scaled model was generated with Simpleware ScanIP^{TM} v4.0, with standard pre-smoothing and mesh refinement settings [47], [48], and will be referred simply as ScanIP Mesh (SM) model. The segmentation pipeline was tested in both noise free and noisy images, corrupted with AWGN with $\sigma_{Noise} = 10$ HU as shown in Fig. 5. Both HD and MSD measures were evaluated to determine the average and maximum error of each segmentation.

Fig. 6 shows the final PC model obtained from the segmentation of Dataset #1, as well its comparison with the Nikon Metris LK V20 gold standard, whereas Fig. 7 depicts the down-scaled segmentation within Simpleware

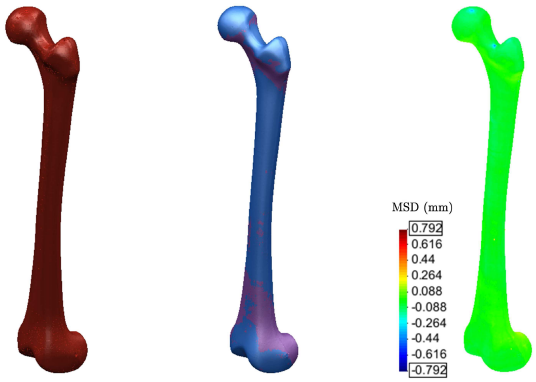


FIGURE 6. Surface mesh obtained with the point cloud model: (a) the point cloud obtained from the segmentation of the phantom femur from dataset #1; (b) the surface mesh generated from the point cloud (magenta) and reference femoral surface obtained with the Nikon Metris LK V20 (blue); and (c) the comparison between the two surfaces using the MSD.

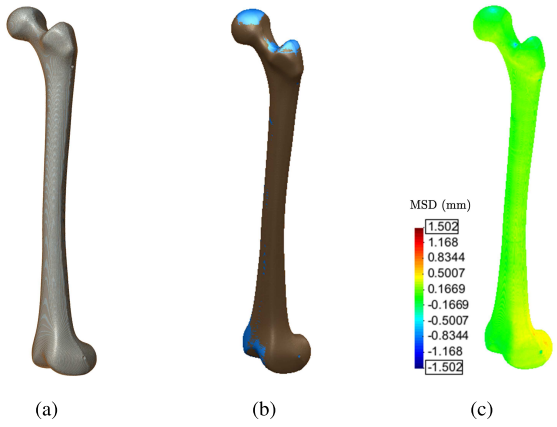


FIGURE 7. Surface mesh obtained with the down scaled model: (a) the domain pre-segmentation and final segmentation after refinement for the dataset #1; (b) the surface mesh generated with the voxelized data obtained after segmentation refinement and reference model obtained with the Nikon Metris LK V20 (blue); and (c) the comparison between the two models using the MSD.

ScanIP^{TM} v4.0 overlapped with the reference surface. A close agreement between both PC and SM models and the gold standard is clearly observed. The agreement between the PC and SM models and the gold standard for all noise-free and noisy Datasets regarding the two distance measures, HD and MSD, is summarized in Table 2 and Table 3, respectively.

A close agreement between the gold-standard and the two surface models was obtained in all Datasets. The average error between the gold-standard surface mesh obtained with Nikon Metris LK V20 and the PC model for DS #1 was 0.077 ± 0.075 mm (Fig. 6), whereas for the SM model the average error was 0.151 ± 0.125 mm (Fig. 7). In the remaining Datasets the maximum average error obtained with the PC model was 0.178 ± 0.170 mm (Table 2 DS #3 + AWGN). Slightly less accurate surface meshes were obtained with the down-scaled model. A maximum average error

TABLE 2. Segmentation accuracy for the PC model according to the dissimilarity measures HD and MSD for the noisy and noise-free Datasets.

	Resolution, (mm)			Model Deviation, (mm)		
Dataset	In-plane		Cross-plane	Average (MSD)	Std.	Maximum (HD)
#1	0.243	0.243	0.30	0.077	0.075	0.792
#2	0.525	0.525	0.30	0.119	0.109	0.841
#3	0.525	0.525	3.00	0.103	0.173	1.843
#4	0.486	0.486	0.50	0.080	0.084	0.851
#5	0.972	0.972	1.00	0.159	0.164	1.073
#1 + AWGN	0.243	0.243	0.30	0.086	0.109	0.971
#2 + AWGN	0.525	0.525	0.30	0.121	0.112	0.908
#3 + AWGN	0.525	0.525	3.00	0.178	0.170	2.187
#4 + AWGN	0.486	0.486	0.50	0.103	0.083	0.907
#5 + AWGN	0.972	0.972	1.00	0.153	0.134	1.199

TABLE 3. Segmentation accuracy for the SM model according to the dissimilarity measures HD and MSD for the noisy and noise-free Datasets.

	Resolution, (mm)			Model Deviation, (mm)		
Dataset	In-plane		Cross-plane	Average (MSD)	Std.	Maximum (HD)
#1	0.243	0.243	0.30	0.151	0.125	0.899
#2	0.525	0.525	0.30	0.348	0.209	1.020
#3	0.525	0.525	3.00	0.262	0.192	1.810
#4	0.486	0.486	0.50	0.284	0.115	1.022
#5	0.972	0.972	1.00	0.617	0.501	2.349
#1 + AWGN	0.243	0.243	0.30	0.172	0.130	1.164
#2 + AWGN	0.525	0.525	0.30	0.368	0.173	1.412
#3 + AWGN	0.525	0.525	3.00	0.342	0.199	1.821
#4 + AWGN	0.486	0.486	0.50	0.273	0.174	1.550
#5 + AWGN	0.972	0.972	1.00	0.572	0.275	1.776

of 0.617 ± 0.501 mm was observed for the SM model (Table 3 DS #5). The results shown of Table 2 and Table 3 allow concluding that the surface meshes produced directly from the high-resolution point cloud provide more accurate descriptions of the ROI for all Datasets. With the PC model the average error is reduced between 51% for DS #1 (from 0.151 ± 0.125 mm to 0.077 ± 0.075 mm) to 26% for DS #5 (from 0.617 ± 0.501 mm to 0.159 ± 0.164 mm) for noise-free images, when compared with the SM model. Similar values were observed for the noisy Datasets, which demonstrates the robustness of the algorithm against image noise.

The PC surface meshes are triangulated from the 3D points, and model smoothness derives directly from the smoothness of the level-set contour. On the contrary, in the SM models the final accuracy is not only dependent on the segmentation accuracy, but also on the settings used for the surface mesh generation. The results show that surface mesh tiling allows an accurate definition of the femoral outer surface, not only in high-resolution scans, but also in scans with more *"clinical"* settings, such as DS #3. Surface meshing directly from the point cloud data effectively avoids the staircase artifacts, commonly observed with larger slice thicknesses, and greatly improves the accuracy of the final model.

In section IV, an average error of 0.073 ± 0.066 mm was obtained for a domain sampling consistent with the FWHM of the PSF. As expected, the image segmentation process is an additional source of model error. Fig. 8 compares graphically the theoretical average and maximum deviation from

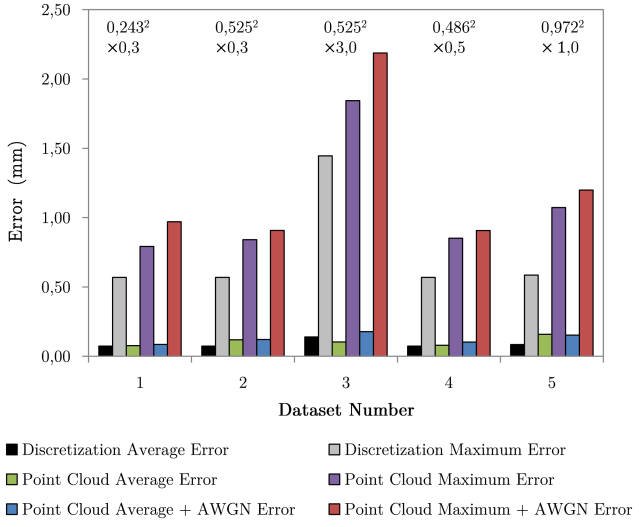


FIGURE 8. Comparison between the domain average (MSD) and maximum (HD) error caused by domain discretization obtained in section IV and the average and maximum error of the final model.

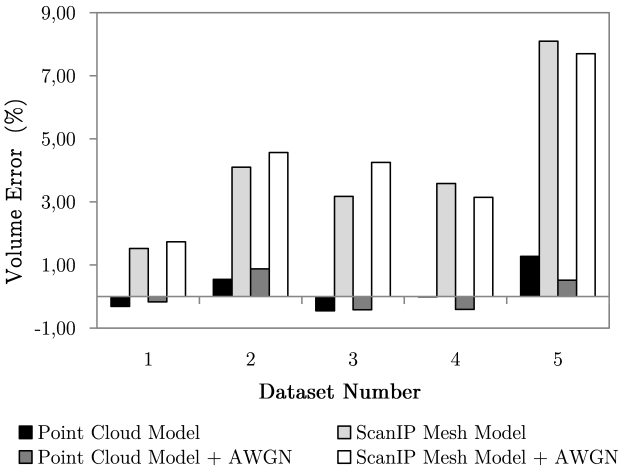


FIGURE 9. Volume dissimilarities between the gold standard model obtained with the Nikon Metris *LK V20*, the point cloud-based model and SM model for noisy and noise-free images.

the gold standard, considering the FWHM of the PSF, and the average and maximum errors for all PC models. For DS #1, the MSD error obtained is in close agreement with the theoretical discretization error: the discretization process accounts for 95.3% and 71.9% of the final average and maximum error, respectively, while the whole chain of image deconvolution, interpolation and segmentation contributes with only 4.7% and 28.1% to the average and maximum errors, respectively. In addition, for the lowest resolution Dataset (DS #5) the discretization process accounts for 53.5% of the final MSD and 54.6% of the final HD errors, with similar values being obtained for the noisy images (Fig. 8). Lastly, a very important statement from Fig. 8: the domain discretization contributes more for the final model deviation than the segmentation refinement chain.

The analysis in terms of volumetric errors is graphically shown in Fig. 9 for all models. The maximum volume deviation was found for the noise free DS #5. The PC models underestimate the ROI in Datasets #1, #3 and #4, more precisely by -0.31%, -0.55% and -0.01%, respectively, and overestimate it in 0.55% for DS #2 and in 1.27% for DS #5. For the down-scaled ScanIP Mesh model (or SM model), volume overestimation ranges from 1.52% in DS #1 to 8.09% in DS #5 for noise free images, and from 1.73% to 7.70% for DS #1 to #5, respectively. Interestingly, for DS #5 the average error and volume deviation decrease with the addition of Gaussian noise. The addition of AWGN to the image caused a slight increase in the average intensity of the foreground, while the average intensity of the background remained almost unchanged. In DS #5 the net effect of adding noise to the image was an overall increase in image contrast. Therefore, the smaller volume error observed in DS #5 + AWGN when compared with DS #5 is caused by the sensitivity of ACWE to image contrast.

VI. DISCUSSION

During image acquisition the CT scanner acts as a low-pass filter, eliminating all the high-spatial frequencies. The sharp transitions between different anatomical regions become unclear, and sometimes barely recognizable. The edge blurring effect often leads to the overestimation of the ROI, and the recovery of the original scene is intrinsically limited by our knowledge about the system's PSF. The ability to precisely determine the PSF is fundamental to extract geometrically accurate models from medical images.

In this work, three different phantom objects are used to determine the CT PSF. Their ability to estimate the system's PSF is compared against the CATPHAN 528, which is the phantom routinely used for quality control. The phantom objects aim to estimate the PSF either directly from the system's impulse response, or indirectly through the estimation of the ESF. Relatively good estimates were obtained with all phantom objects. However, the brass wire with 0.10 mm of diameter provided the most accurate estimate of the PSF with a standard deviation error of only 2.9%, when compared with the PSF obtained with the CATPHAN 528. Similar values for the Toshiba Aquilion^{TM} 64 system were obtained in [49].

Unlike direct measurements, the estimation of the PSF through the analysis of the system's response to strong edges implies additional signal processing steps. Gaussian filtering has been associated with edge bias by several authors [34], [35]. In addition, the ESF may be affected by the presence of image noise during the calculation of the gradient of the attenuation profile. Noise must be eliminated prior to the calculation of the gradient. In this context, Mori and Machida [39] proposed an additional blurring correction step to eliminate the Gaussian smoothing effect from the estimate of the MTF estimate. Instead of correcting the result, Wavelet thresholding was used in this work to avoid linear signal filtering. The results show that although less accurate than the wire phantom, the estimate of the

PSF through the ESF and Wavelet denoising produces quite good approximations to the real CT spatial blurring function. The Wavelet thresholding provides the means to effectively reduce the noise level along the 1D attenuation profile, and to accurately estimate the ESF without the need for any compensation procedure.

In practice, the accurate quantification of the system's PSF allow us to accurately recover the original image using more standard image deconvolution algorithms. In fact, the accurate estimate of this quantity proved to be quite important to obtain accurate segmentations from image data. In [23] it was concluded that reversing the blurring degradation, prior to segmentation, is essential to construct accurate FE models from medical imaging.

Regarding image segmentation, a new two-step segmentation pipeline was proposed. The segmentation outcomes were validated through a phantom study, where the final models were compared against a gold-standard surface mesh. The gold-standard representation of the phantom object was acquired with a coordinate measuring machine Nikon Metris LK V20 with a digital line scanner LC60-D, which guarantees 28 μm of spatial accuracy.

For all Datasets, a maximum MSD of 0.178 mm from the gold standard model was obtained. The maximum HD was 2.187 mm, for DS #3 corrupted with AWGN (Table 2). The accurate estimation of the system's PSF, together with image interpolation and level-set segmentation provide quite good segmentations of the target geometry. In addition, the results show that surface meshes computed directly from high-resolution 3D point clouds provide more accurate representations of the phantom bone, when compared with the standard surface meshing pipeline from the voxelized data.

The surface models computed from the high-resolution point cloud data are smooth and relatively independent from the surface meshing. Surface triangulation from adjacent contours (or surface tiling) also avoids staircase artifacts, due to large section variation between consecutive slices, caused by large slice thicknesses [50]. The advantage of using surface tiling is observable when comparing the results obtained in datasets #2 and #3, in which the obtained MSD differs by only 0.016 mm. Nevertheless, the largest maximum deviation observed in the PC models may also be a consequence of surface mesh tiling. With surface mesh tiling, high-accuracy models may be extracted from the image data, even when images are acquired with more clinical settings. This may help reducing patient exposure to the radiation, without a significant loss in model accuracy.

The results also show that segmentation accuracy is more dependent on the reconstruction FOV than on the slice thickness. This observation in in agreement with observations found in other phantom studies [13], [45]. In addition and unlike previous works [24], our results also show that bone segmentation through the level-set method is accurate and possesses numerous advantages against other segmentation methods. The level-set method is robust to noise, it guarantees contour smoothness, is topologically flexible, and allows a straightforward incorporation of user-defined spatial constraints.

Sampling the gold-standard model with a voxel size of FWHM $=$ 0.88 mm produces an MSD of approximately 8.3% of the sampling (0.073 ± 0.066 mm). Accordingly, the discretisation process (image acquisition and reconstruction) is the main limiting factor of the segmentation pipeline. The average error due to domain sampling ranges from 95.31% for DS #1 to 53.46% for DS #5 of the final PC model error. Therefore, the discretization MSD between the two models (reference and segmented) always accounts for more than 50% of the final average error (Fig. 8). The FWHM of PSF may also explain the average errors obtained with dataset #1 to #4, which are very similar despite of voxel sizes, and the considerable increase in the average error observed for DS #5. In dataset #1 to #4, the FWHM of the PSF is the main limitation to the spatial accuracy of the model. However, in DS #5 the voxel size is greater than the FWHM of the PSF and becomes the main limiting factor.

The average errors obtained show that in the whole chain of image deconvolution, image up-scaling through cubic spline interpolation, and subsequent segmentation, there is almost no information degradation or distortion, due to the proposed image processing pipeline. In fact, this is particularly evident in DS #1, where the whole image processing chain accounts for only 4.7% of the final error (for noise-free images), and 14.6% for images corrupted with AWGN. Our observations also corroborate the observations of [46], which concluded that the domain discretization contributes in a larger extent to the final error of the surface mesh model.

TABLE 4. Bone segmentation accuracy obtained in other studies found in the literature.

	Resolution (mm)			Model Deviation (mm)		
	In-plane		Cross-plane	Average (MSD)	Std.	Maximum (HD)
[20]	0.312	0.312	2.00	0.03	0.52	5.75
[12]	0.200	0.200	0.50	0.20	0.09-0.13	0.26
[51]	0.293	0.293	0.63	0.46	0.03	0.49
[18]	0.195	0.195	1.00	0.21	0.12	0.47
[52]	0.480	0.480	0.50	0.14	—	1.81
[22]	0.589	0.589	1.00	0.02	0.77	—
[21]	0.390	0.390	0.50	0.18	0.20	—
[23]	0.480	0.480	1.00	0.14	0.11	—

The segmentation in two steps proposed in this work is very similar to other methods found in the literature. In Table 4, the main results obtained in other studies are summarized. The data is presented according to the original papers. Therefore, it is important to keep in mind that the results might have been obtained in different conditions than the ones here described. There are three studies that have an average error smaller than the maximum average error, obtained with the proposed segmentation refinement protocol, namely the optimized 600 HU threshold proposed in [20], the Levenberg-Marquardt-based algorithm proposed in [22], and the gradient descent algorithm proposed in [23]. Regarding the optimized 600 HU threshold in [20], the average error ranged from -0.20 mm to 0.20 mm in the

eight femurs considered. The maximum average error is larger than the maximum average error obtained with the new algorithm proposed, for all Datasets. Even in lower resolution scans (e.g. DS #3 and #5), the proposed protocol produces more accurate results (see Table 2). Furthermore, the standard deviation of 0.77 mm found in [22] shows that the algorithm mainly oscillate around the true surface of the cortical bone. In [23] slightly less oscillating results were obtained. However, for a comparable spatial resolution (DS #4), the results in the current study have a smaller average error of 0.080 mm (with 68.2% of the surface points within the interval of -0.004 to 0.164 mm) for noise free data, and with an average error of 0.103 mm (with 68.2% of the surface points within the interval of 0.020 to 0.185 mm) for noisy data.

The practical implications of an average deviation from the gold standard of 0.178 mm in custom implant modelling are still unclear. In an early study, Carlsson et al. [53] found that implant-bone surface gaps of 0.35 mm or more were not bridged by cortical bone, and that this value is close to the critical gap width for which direct lamellar bone apposition occurs onto unloaded implants. In an animal study, Pazzaglia et al. [54] found that there was no bone integration at the interface of roughened titanium rods with 0.30 mm of diametrical gaps at the bone-implant interface. The average error obtained may have minor practical implications in the performance of the custom implants developed from the 3D models for both orthopaedic and trauma applications. In [20] an inward offset of 0.50 mm to the contours obtained with the 600 HU contour was proposed to avoid custom femoral stem over dimensioning. Endosteal contour shrinkage is needed due to the large standard deviation (0.52 mm) around the average error (see Table 4). Traditional surgical techniques may play a more relevant role in the implant outcome than the geometrical accuracy obtained during implant modelling (see for instance [55] and [56]).

The main limitations of the present study are directly linked with the limitations often associated with every phantom study. Phantom studies are flexible to parameterize and allow the simulation of different acquisition and reconstruction protocols and noise levels, among others. In this type of studies, the ground truth may be accurately characterized, and does not change under different environmental conditions. Nevertheless, phantom studies may not be realistic enough to model the complexity of the real data. The CT images used in this analysis have good contrast and are almost noise free. Such images are very difficult or even impossible to obtain in practice. The ACWE are very robust to noise but sensitive to image contrast, therefore in real data the performance of the present algorithm may be slightly degraded. Furthermore, although the use of ACWE enables some degree of flexibility in the pre-segmentation step. There is still, however, some dependency on the curve initialization. Image under-segmentation seems to be more problematic than over-segmentation, hence in ambiguous pixels it is better to over-segment the domain rather than excluding the pixels. The elimination of such a dependency and the extension of the protocol to the third dimension, in order to obtain a 3D point cloud instead of the current 2D contours are natural evolutions to the current algorithm. These may facilitate the pre-segmentation step and may help to overcome the limitations of surface mesh tiling near bone ends, and allow the production of more accurate models near articulating surfaces. Another open issue is the accuracy of the segmentation along the endosteal surface of the cortical bone. A high resolution scan (such as micro-CT) may be needed to further validate the performance of the algorithm along the inner surface of the bone. These issues will be considered in future work.

VII. CONCLUSIONS

Currently, CT imaging is seen as the most successful medical imaging technique. It has widespread across all medical fields, and it is oftenly used as a complementary tool for diagnosis, as well as in the development of patient-specific biomechanical and finite element models. One of the most fundamental step in custom implant development is the accurate description of the Human skeletal system. Therefore, it is important to develop segmentation protocols that guarantee segmentation accuracy, especially in a custom implant development and manufacturing pipeline.

In this work a new two-step segmentation pipeline is proposed for accurate bone segmentation from CT image data. The proposed methodology handles segmentation variability by allowing a first free pre-segmentation step, where the user can employ the necessary means to obtain an approximation to the target segmentation. The second step is standardized, fully automatic and encompasses image restoration, cropping, interpolation, and level-set segmentation. The proposed methodology produces accurate estimates of the target geometries with a maximum average deviation of 0.178 mm. Results show that surface meshes extracted directly from the high-resolution point cloud describe more accurately the target ROI. The accuracy of the point cloud model is mainly affected by the image acquisition and reconstruction, rather than by the image segmentation and surface meshing processes. In addition, the CT machine PSF can be accurately determined using a brass alloy wire phantom with 0.10 mm of diameter. The direct measures provide slightly more accurate estimates of the system's Point Spread Function, when compared with indirect measures based on the Edge Spread Function.

The proposed segmentation protocol shows some dependence on curve initialization, and on image contrast. The incorporation of the attenuation profiles normal to the level-set curve may avoid this dependency, driving the curve to the correct position along the bone's boundary. In addition, point cloud models fail to produce smooth meshes near the articular surfaces, due to the surface mesh tiling process. The extension of the refinement pipeline to a truly 3D procedure may probably avoid this shortcoming. However, extending the whole refinement step to 3D, without increasing its computational complexity is challenging.

ACKNOWLEDGMENT

The authors also would like to acknowledge Hospital CUF, Porto (Portugal), Clínica Dr. Campos Costa, Porto (Portugal), and ISQ, Instituto de Soldadura e Qualidade for all technical support provided during this work.

REFERENCES

[1] A. E. Anderson, C. L. Peters, B. D. Tuttle, and J. A. Weiss, "Subject-specific finite element model of the pelvis: Development, validation and sensitivity studies," *J. Biomech. Eng.*, vol. 127, no. 3, pp. 364–373, 2005.

[2] L. Antiga, M. Piccinelli, L. Botti, B. Ene-Iordache, A. Remuzzi, and D. A. Steinman, "An image-based modeling framework for patient-specific computational hemodynamics," *Med. Biol. Eng. Comput.*, vol. 46, no. 11, pp. 1097–1112, 2008.

[3] O. L. A. Harrysson, Y. A. Hosni, and J. F. Nayfeh, "Custom-designed orthopedic implants evaluated using finite element analysis of patient-specific computed tomography data: Femoral-component case study," *BMC Musculoskeletal Disorders*, vol. 8, no. 1, p. 91, 2007.

[4] A. V. Lombardi, Jr., K. R. Berend, and J. B. Adams, "Patient-specific approach in total knee arthroplasty," *Orthopedics*, vol. 31, no. 9, pp. 927–930, 2008.

[5] F. Gelaude, J. Vander Sloten, and B. Lauwers, "Accuracy assessment of CT-based outer surface femur meshes," *Comput. Aided Surgery*, vol. 13, no. 4, pp. 188–199, 2008.

[6] B. A. Klatt, N. Goyal, M. S. Austin, and W. J. Hozack, "Custom-fit total knee arthroplasty (OtisKnee) results in malalignment," *J. Arthroplasty*, vol. 23, no. 1, pp. 26–29, 2008.

[7] P. S. Walker and D. D. Robertson, "Design and fabrication of cementless hip stems," *Clin. Orthopaedics Rel. Res.*, vol. 235, pp. 25–34, Oct. 1988.

[8] J. A. Mandell, D. R. Carter, S. B. Goodman, D. J. Schurman, and G. S. Beaupré, "A conical-collared intramedullary stem can improve stress transfer and limit micromotion," *Clin. Biomech.*, vol. 19, no. 7, pp. 695–703, 2004.

[9] A. F. Mavrogenis, R. Dimitriou, J. Parvizi, and G. C. Babis, "Biology of implant osseointegration," *J. Musculoskeletal Neuronal Interact.*, vol. 9, no. 2, pp. 61–71, 2009.

[10] T. N. Hangartner and V. Gilsanz, "Evaluation of cortical bone by computed tomography," *J. Bone Mineral Res.*, vol. 11, no. 10, pp. 1518–1525, 1996.

[11] S. Prevrhal, K. Engelke, and W. A. Kalender, "Accuracy limits for the determination of cortical width and density: The influence of object size and CT imaging parameters," *Phys. Med. Biol.*, vol. 44, no. 3, pp. 751–764, 1999.

[12] Y. Kang, K. Engelke, and W. A. Kalender, "A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 586–598, May 2003.

[13] M. Ohkubo, S. Wada, M. Kunii, T. Matsumoto, and K. Nishizawa, "Imaging of small spherical structures in CT: Simulation study using measured point spread function," *Med. Biol. Eng. Comput.*, vol. 46, no. 3, pp. 273–282, 2008.

[14] G. Dougherty, *Digital Image Processing for Medical Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2009.

[15] A. Maloul, J. Fialkov, and C. Whyne, "The impact of voxel size-based inaccuracies on the mechanical behavior of thin bone structures," *Ann. Biomed. Eng.*, vol. 39, no. 3, pp. 1092–1100, 2011.

[16] L. W. Goldman, "Principles of CT: Radiation dose and image quality," *J. Nucl. Med. Technol.*, vol. 35, no. 4, pp. 213–225, 2007.

[17] C. F. Hildebolt, M. W. Vannier, and R. H. Knapp, "Validation study of skull three-dimensional computerized tomography measurements," *Amer. J. Phys. Anthropol.*, vol. 82, no. 3, pp. 283–294, 1990.

[18] J. Wang, M. Ye, Z. Liu, and C. Wang, "Precision of cortical bone reconstruction based on 3D CT scans," *Comput. Med. Imag. Graph.*, vol. 33, no. 3, pp. 235–241, 2009.

[19] P. Rothuizen, L. van Erning, and R. Huiskes, "The accuracy of criteria for automatic 3-D graphics reconstruction of bone from computer tomography," in *Biomechanics: Basic and Applied Research*. New York, NY, USA: Springer-Verlag, 1987, pp. 109–114.

[20] A. Aamodt *et al.*, "Determination of the Hounsfield value for CT-based design of custom femoral stems," *J. Bone Joint Surgery (Brit.)*, vol. 81-B, no. 1, pp. 143–147, 1999.

[21] K. Rathnayaka, T. Sahama, M. A. Schuetz, and B. Schmutz, "Effects of CT image segmentation methods on the accuracy of long bone 3D reconstructions," *Med. Eng. Phys.*, vol. 33, no. 2, pp. 226–233, 2011.

[22] G. M. Treece, A. H. Gee, P. M. Mayhew, and K. E. S. Poole, "High resolution cortical bone thickness measurement from clinical CT data," *Med. Image Anal.*, vol. 14, no. 3, pp. 276–290, 2010.

[23] A. Pakdel, N. Robert, J. Fialkov, A. Maloul, and C. Whyne, "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.*, vol. 57, no. 23, pp. 8099–8116, 2012.

[24] W. Yao, P. Abolmaesumi, M. Greenspan, and R. E. Ellis, "An estimation/correction algorithm for detecting bone edges in CT images," *IEEE Trans. Med. Imag.*, vol. 24, no. 8, pp. 997–1010, Aug. 2005.

[25] B. Preim and D. Bartz, *Visualization in Medicine: Theory, Algorithms, and Applications* (The Morgan Kaufmann Series in Computer Graphics). Amsterdam, The Netherlands: Elsevier, 2007.

[26] D. J. Goodenough, *Catphan 500 and 600 Manual: Multi-Slice CT*. Greenwich, NY, USA: The Phantom Laboratory, 2014.

[27] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Trans. Med. Imag.*, vol. 19, no. 7, pp. 739–758, Jul. 2000.

[28] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[29] L. A. Vese and T. F. Chan, "A multiphase level set framework for image segmentation using the Mumford and Shah model," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 271–293, 2002.

[30] S. Lankton, "Sparse field methods," Georgia Inst. Technol., Atlanta, GA, USA, Tech. Rep., 2009.

[31] P. Campisi and K. Egiazarian, Eds., *Blind Image Deconvolution: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2007.

[32] C. Solomon and T. Breckon, *Fundamentals of Digital Image Processing: A Practical Approach With Examples in MATLAB*. New York, NY, USA: Wiley, 2011.

[33] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 7, pp. 629–639, Jul. 1990.

[34] P. R. S. Mendonça, D. Padfield, J. Miller, and M. Turek, "Bias in the localization of curved edges," in *Computer Vision*. Berlin, Germany: Springer-Verlag, 2004, pp. 554–565.

[35] H. Bouma, A. Vilanova, L. J. Van Vliet, and F. A. Gerritsen, "Correction for the dislocation of curved surfaces caused by the PSF in 2D and 3D CT images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 9, pp. 1501–1507, Sep. 2005.

[36] R. L. Easton, Jr., *Fourier Methods in Imaging*. New York, NY, USA: Wiley, 2010.

[37] S. Dore and Y. Goussard, "Experimental determination of CT point spread function anisotropy and shift-variance," in *Proc. 19th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, vol. 2. Oct./Nov. 1997, pp. 788–791.

[38] E. Samei, M. J. Flynn, and D. A. Reimann, "A method for measuring the presampled MTF of digital radiographic systems using an edge test device," *Med. Phys.*, vol. 25, no. 1, pp. 102–113, 1998.

[39] I. Mori and Y. Machida, "Deriving the modulation transfer function of CT from extremely noisy edge profiles," *Radiol. Phys. Technol.*, vol. 2, no. 1, pp. 22–32, 2009.

[40] D. L. Donoho and I. M. Johnstone, "Adapting to unknown smoothness via wavelet shrinkage," *J. Amer. Statist. Assoc.*, vol. 90, no. 432, pp. 1200–1224, 1995.

[41] M. Pinheiro, F. Martins, and J. L. Alves, "Towards a medical image enhancement procedure based on wavelet transform analysis," *J. Biomech.*, vol. 45, p. S462, Jul. 2012.

[42] N. Joshi, R. Szeliski, and D. Kriegman, "PSF estimation using sharp edge prediction," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2008, pp. 1–8.

[43] E. Pantin, J.-L. Starck, and F. Murtagh, "Deconvolution and blind deconvolution in astronomy," in *Blind Image Deconvolution: Theory and Applications*. Boca Raton, FL, USA: CRC Press, 2007, pp. 100–138.

[44] T. K. Goto *et al.*, "The accuracy of 3-dimensional magnetic resonance 3D vibe images of the mandible: An in vitro comparison of magnetic resonance imaging and computed tomography," *Oral Surgery, Oral Med., Oral Pathol., Oral Radiol., Endodontol.*, vol. 103, no. 4, pp. 550–559, 2007.

[45] M. I. Meurer *et al.*, "Influence of tomographic slice thickness and field of view variation on the reproduction of thin bone structures for rapid prototyping purposes—An *in vitro* study," *Open J. Radiol.*, vol. 3, no. 1, pp. 12–25, 2013.

[46] S. Cortez, J. C. P. Claro, and J. L. Alves, "3D reconstruction of a spinal motion segment from 2D medical images: Objective quantification of the geometric accuracy of the FE mesh generation procedure," in *Proc. IEEE 3rd Portuguese Meeting Bioeng. (ENBENG)*, Feb. 2013, pp. 1–6.
[47] *ScanIP+FE+CAD Reference Guide*, Simpleware, Exeter, U.K., 2011.
[48] P. G. Young, T. B. H. Beresford-West, S. R. L. Coward, B. Notarberardino, B. Walker, and A. Abdul-Aziz, "An efficient approach to converting three-dimensional image data into highly accurate computational models," *Philos. Trans. Roy. Soc. London A, Math. Phys. Sci.*, vol. 366, no. 1878, pp. 3155–3173, 2008.
[49] B. M. Verbist, R. M. S. Joemai, W. M. Teeuwisse, W. J. H. Veldkamp, J. Geleijns, and J. H. M. Frijns, "Evaluation of 4 multisection CT systems in postoperative imaging of a cochlear implant: A human cadaver and phantom study," *Amer. J. Neuroradiol.*, vol. 29, no. 7, pp. 1382–1388, 2008.
[50] I. Braude, J. Marker, K. Museth, J. Nissanov, and D. Breen, "Contour-based surface reconstruction using MPU implicit models," *Graph. Models*, vol. 69, no. 2, pp. 139–157, 2007.
[51] K. Oka, T. Murase, H. Moritomo, A. Goto, K. Sugamoto, and H. Yoshikawa, "Accuracy analysis of three-dimensional bone surface models of the forearm constructed from multidetector computed tomography data," *Int. J. Med. Robot. Comput. Assist. Surgery*, vol. 5, no. 4, pp. 452–457, 2009.
[52] X. Liang *et al.*, "A comparative evaluation of cone beam computed tomography (CBCT) and multi-slice CT (MSCT). Part II: On 3D model accuracy," *Eur. J. Radiol.*, vol. 75, no. 2, pp. 270–274, 2010.
[53] L. Carlsson, T. Röstlund, B. Albrektsson, and T. Albrektsson, "Implant fixation improved by close fit: Cylindrical implant—Bone interface studied in rabbits," *Acta Orthopaedica*, vol. 59, no. 3, pp. 272–275, 1988.
[54] U. E. Pazzaglia, F. Brossa, G. Zatti, R. Chiesa, and L. Andrini, "The relevance of hydroxyapatite and spongious titanium coatings in fixation of cementless stems," *Arch. Orthopaedic Trauma Surgery*, vol. 117, nos. 4–5, pp. 279–285, 1998.
[55] H. A. Paul *et al.*, "Development of a surgical robot for cementless total hip arthroplasty," *Clin. Orthopaedics Rel. Res.*, vol. 285, pp. 57–66, Dec. 1992.
[56] A. Manzotti, P. Cerveri, E. De Momi, C. Pullen, and N. Confalonieri, "Does computer-assisted surgery benefit leg length restoration in total hip replacement? Navigation versus conventional freehand," *Int. Orthopaedics*, vol. 35, no. 1, pp. 19–24, 2011.

MANUEL PINHEIRO was born in Guimarães, Portugal, in 1986. He received the M.S. degree in biomedical engineering from the University of Minho, Guimarães, in 2009, where he is currently pursuing the Ph.D. degree in biomedical engineering.

He became a Post-Doctoral Researcher with the Medical and Biological Engineering Research Group, University of Hull, Hull, U.K., in 2015. His main research interests are medical image processing, image segmentation, custom implant development, the biomechanics of the human musculoskeletal system, and finite element analysis.

J. L. ALVES was born in Braga, Portugal, in 1971. He received the M.S. degree in mechanical engineering and the Ph.D. degree in computational mechanics and manufacturing technologies from the University of Minho, Portugal, in 1999 and 2003, respectively.

He was an Invited Researcher with the Physical and Chemical Institute, RIKEN, Japan, from 2007 to 2011, where he collaborated in the framework of V-CAD project. He has been an Assistant Professor with the Department of Mechanical Engineering, School of Engineering, University of Minho, since 2003. Since 2013, he has been an Adjunct Professor with the REEF/University of Florida, Shalimar, USA. He has authored or co-authored several finite element solvers, such as V-Biomech, V-MultiMat, and DD3IMP, and holds three patents. His research interests include constitutive modeling, elasticity and plasticity, finite elements, imaging processing, finite element mesh generation and optimization, biomechanics, and the constitutive modeling of hard and soft tissues.

He received the Esaform Scientific Prize in 2006, which intends to distinguish a young confirmed scientist, who has brought an outstanding contribution in the field of materials forming.