Uncertainty Reduction in Contour-Based 3D/2D Registration of Bone Surfaces

Xolisile O. Thusini[1(✉)], Cornelius J. F. Reyneke[1], Jonathan Aellen[3],
Andreas Forster[3], Jean-Rassaire Fouefack[1,2], Nicolas H. Nbonsou Tegang[1],
Thomas Vetter[3], Tania S. Douglas[1], and Tinashe E. M. Mutsvangwa[1,3]

[1] Division of Biomedical Engineering, University of Cape Town,
Cape Town, South Africa
`thsxol001@myuct.ac.za`
[2] Department of Image and Information Processing, IMT-Atlantique, Brest, France
[3] Department of Mathematics and Computer Science, University of Basel, Basel,
Switzerland

Abstract. The reconstruction of $3D$ bone shape from $2D$ X-ray contours is an ill-posed problem. For medical applications, it is important to estimate the uncertainty of the reconstructions. While traditional optimisation methods produce a single point-estimate, we frame the problem as Bayesian inference. We apply a Monte Carlo sampling based non-rigid $3D$ to $2D$ registration recovering the posterior distribution of plausible reconstructions. This provides insight into the uncertainty of the inferred $3D$ reconstruction. As an application, we demonstrate the use of the method in selecting X-ray viewing conditions in order to maximise accuracy while minimising reconstruction uncertainty. We evaluated reconstruction accuracy and variance for the femur bone from bi-planar views.

Keywords: 3D reconstruction · Statistical shape model · $3D/2D$ registration · Bayesian inference · MCMC sampling · Optimal viewing-angle selection

1 Introduction

Non-rigid three-dimensional-to-two-dimensional ($3D/2D$) registration may be used to obtain a patient-specific $3D$ surface reconstruction of a bone from a limited number of $2D$ patient X-ray images. Such a reconstruction is performed by iteratively adjusting the parameters of a $3D$ statistical shape model (SSM), based on the information inferred from the X-ray views. These surface reconstructions can then be used to aid in clinical tasks such as surgical planning and post-operative evaluation while avoiding the higher costs or relatively greater patient irradiation associated with some $3D$ modalities (computed tomography). The main clinical application area of such reconstructions is within orthopaedics. In order to reduce exposure to ionising radiation, and reduce computation time, the number of X-ray views is often limited to one or two (Hurvitz and Joskowicz

2008). The viewing angle(s) are selected, either by adjusting the patient or by adjusting the C-arm of the X-ray machine (if one is available). In contour-based reconstruction, the chosen viewing angle(s), together with the type of bone, have an influence on the points that can reliably be identified, and thus also influence the uncertainty of the estimated patient-specific $3D$ bone surface. A decision, therefore, has to be made about which X-ray view(s) to use as many $3D$ shapes can account for the same X-ray image projection. Additionally, since different types of symmetry exist for different types of bone (femur vs. pelvis), the viewing angles that provide the least ambiguity - and maximise accuracy - are not always obvious (Suter et al. 2020). While the femur bone is relatively simple, and a 90° bi-planar viewing angle seems intuitive (since a maximum decorrelation of information between the two views in ensured) (Fleute and Lavallée 1999; Baka et al. 2011), bones such as the scapula or pelvis exhibit a profound increase in structural superposition for certain viewing angles. In their work, Sadowsky et al. (2007) opted to use a 90° and 45° separation for the pelvis (Sadowsky et al. 2007). It is evident that these methods rely mostly on user intuition, or heuristics, which lead to reconstruction error and/or inconsistent performance. Furthermore, maximum-point-estimate methods are employed, where the effect of different sources of uncertainty (such as $2D$ projection) are not investigated. Markov Chain Monte Carlo (MCMC) methods for image analysis are used to assess the uncertainty and have been reported to solve inference problems in object recognition, including the parsing of natural scene images, face analysis and segmentation (Zhu et al. 2000; Tu et al. 2005; Schönborn et al. 2017; and Egger 2017). The probabilistic nature of MCMC approaches allows reason about uncertainty of the image interpretation while simultaneously performing image registration.

Here we propose a framework for contour-based $3D$-from-2D reconstruction. We frame the problem as model fitting based on Bayesian inference. We estimate the full posterior distribution of model instances given the image, and not only a single maximum-point-estimate as Sadowsky et al. (2007). The posterior has no closed form solution, and thus we resort to MCMC based approaches. We use the Metropolis-Hastings algorithm in combination with a given shape prior model considering minor pose variation. The inference is driven by a propose-and-verify architecture. A new sample is proposed from a proposal distribution given the current state. Then the sample is either accepted or rejected based on a acceptance criterion as originating from the posterior distribution over model parameters.

Since the exploration of the space with random proposals is slow at high dimensions and many samples are not independent, we propose the use of Hamiltonian Monte Carlo (HMC) as the second approach (Duane et al. 1987; Neal et al. 2011) to explore the search space more efficiently, while retaining accurate posteriors. We demonstrate how to use the inferred posterior distribution to compute the uncertainty of the reconstructed surface. Furthermore, we propose a method for automatic contour selection (on the $3D$ reference surface), capable of simulating the view-dependent contour, which can help determine how

informative a given contour may be. By combining our algorithm with a contour generation method, we provide a means of quantifying uncertainty reduction for a particular X-ray view (or views). The efficacy of the framework is evaluated through a series of experiments using simulated images. Our contributions are as follows: 1) an MCMC-sampling method for $3D$-from-$2D$ bone reconstruction using the random walk Metropolis-Hastings method; 2) an improved sampling algorithm using the HMC method; 3) a method for estimating the uncertainty of a $3D$ reconstruction; and 4) an application of the proposed framework to determine optimal bi-planar X-ray viewing-angles.

2   Methods

2.1   Statistical Shape Models

We employ the method described by Lüthi et al. (2017) to train an SSM of the bone-of-interest. Building the SSM basically includes the extraction of the mean shape and a number of modes of variation across the individuals in the training dataset of typical shapes $\{\Gamma_n, n = 1, ..., N\}$ segmented from the CT scans of the population. The first step in computing the mean shape is to extract shape features from each training sample and align all the samples with respect to the reference shape using these features. The second step is the reduction of the training dataset into a small set of modes that best represent the observed shape variations. This is accomplished using principal component analysis (PCA) (Hotelling 1933). The PCA models only represent shapes that are in the linear span of the given training examples. The drawback of this specificity is that not all the target shapes can be fully represented in the model. To capture the complete PCA model space, a lot of training samples are needed which in reality is not a simple task. To overcome this, the work by Lüthi et al. (2017) extended the PCA models to include modeling of deformations using Gaussian process (GP). Hence the resulting model is the probability distribution defined on the deformations. Including the GP means that the deformations are modelled with Gaussian function $\mathcal{GP}(\mu, k)$, where μ and k are mean and covariance functions, respectively. In a Bayesian inferencing framework, the model is used as a prior distribution $P(\theta)$, where θ denotes the model parameters that are optimised during contour-fitting.

2.2   Synthetic Contour Generation, Projection and Detection

We demonstrate the validity of our approach in a simplified setting. To generate synthetic 2D images of contours, we produce a $2D$ silhouette from the 3D mesh. Our approach requires that we know the silhouette vertex points.

Our algorithm consists of two steps. First, we render a mesh as a binary image. Second, we find all vertices which account for the silhouette.

Let $M(\theta)$ be the mesh described by the model parameters θ. Next we render a binary occupancy image o_θ with the value $o_\theta(p) = 1$ when the object is visible

at the pixel location p otherwise 0. Let N_p denote the set of adjacent pixels of pixel p, e denotes the eye vector, p_v the pixel location of vertex v in the binary image, S_v the set of adjacent surfaces of v with n_s denoting the normal of the surface $s \in S_v$. Then the collection of silhouette points C is defined as:

$$C = \{ \ v \mid v \in M(\theta), \ \exists s', s'' \in S_v, \ s' \neq s'' \text{ with } sgn(n_{s'} \cdot e) \neq sgn(n_{s''} \cdot e) \\ \text{and } \exists p' \in N_p \text{ with } o_\theta(p') = 0 \} \ . \tag{1}$$

The vertex set C in Eq. 1 contains all points which are on the silhouette and have adjacent surfaces which face in a different direction when projected onto the eye vector. Even though we generate a $2D$ silhouette image given a $3D$ shape, we still calculate distances using only the $2D$ silhouette.

This synthetic contour generation process is also used during the optimisation process in the model fitting.

2.3   Estimation of the Posterior Distribution

In order to estimate the posterior distribution $P(\theta|C)$ of the SSM, when provided with a set of pre-selected view-dependent $2D$ contours as observations, we replace a straight-forward optimisation approach with a MCMC-based inference algorithm. The inferred solution is a distribution, rather than a single maximum estimate.

Random Proposals. We adopt the MCMC approach of Schönborn et al. (2017), where the inference is driven by the proposed-verification architecture of the Metropolis-Hasting (MH) algorithm, as our first sampling approach. This algorithm draws random samples θ' from a proposal distribution $Q(\theta'|\theta)$ and transforms them into samples stemming from a target distribution $P(\theta|C)$ by accepting a proposed sample $\theta_{t+1} \to \theta'$ with the following probability:

$$a = \min\{\frac{P(\theta'|C)}{P(\theta|C)} \frac{Q(\theta|\theta')}{Q(\theta'|\theta)}, 1\}. \tag{2}$$

The algorithm may otherwise reject the sample and keep the current one $\theta_{t+1} \to \theta$. The function Q generates proposals θ' that are possible model parameter updates. Our implementation of the MCMC sampling-based registration algorithm integrates prior knowledge and contour observations using accept-reject filters. In this setting, we initialise the Markov chain as the model mean which later is updated after each accepted sample. Each accepted sample corresponds to a set of model parameters which provide an estimated bone surface. The resultant surface meshes are then used to represent the estimated posterior distribution for a specific set of contour points (corresponding to a specific X-ray view).

Guided Proposals. We use the HMC approach to efficiently explore the complex target distribution and obtain a more accurate variance estimator. The computational expense to create independent samples for a d-dimensional space generally scales with $d^{5/4}$, while it is d^2 for a random walk MH method. To define the Hamiltonian dynamics (Neal et al. 2011; Zoppo 2018), which create proposals which are nearly independent from the current state but still have a high acceptance rate, we extend a d-dimensional current state q with d-dimensional momentum variables p. The Hamiltonian function $H(q, p)$ is the sum of the potential energy $U(q)$ and the kinetic energy $K(p)$ formulated as:

$$H(q,p) = U(q) + K(p), \tag{3}$$

$$U(q) = -\log\left[P(q)l(q|C)\right] \text{ and } K(p) = p^T M^{-1} p/2.$$

Here the quadratic form of the kinetic energy is used to model the Gaussian targets. The mass matrix M is a free parameter for linear transformation. We use our estimate of the covariance matrix for M. The posterior distribution takes the role of the position variable, q which is expressed using the potential energy with $P(q)$ being the prior, and $l(q|C)$ is the likelihood.

Algorithm 1 explains how Eq. 3 is approximated for a given set of initial conditions, using a step-size ϵ, and number of leapfrog steps L. The dynamics are randomly perturbed by the momentum values drawn from the normal distribution to ensure that the space is explored at each sampling iteration.

Algorithm 1: Hamiltonean Monte Carlo.

HMC = f(U,gradU,ϵ,L,current_q)
Initialise
q= current_q
M $\leftarrow$ diagonal scale matrix
p $\leftarrow$ N(0,M)
current_p=p
p = p - ϵ * gradU(q)/2; % ensure volume preservation and reversibility
for $i=1,...,L$ **do**
$\quad$ q= q + ϵ * p
$\quad$ **if** $i \mathrel{!}= L$ **then**
$\quad\quad$ | p = p - ϵ * gradU(q)
$\quad$ **end**
end
p = p - ϵ * gradU(q)/2
p = -p
if $N(0,1) \leq exp\left[U(current\_q) - U(q)|\frac{||current\_p||^2}{2} - \frac{||p||^2}{2}\right]$ **then**
$\quad$ | **return** q
end
else
$\quad$ | **return** current_q
end

Contour Likelihoods. We compute the posterior of the shape reconstruction $P(\theta|C)$ from a prior $P(\theta)$ and a contour likelihood $l(\theta; C)$

$$P(\theta|C) \propto l(\theta; C)P(\theta). \tag{4}$$

The likelihood functions are important in the acceptance step of both MC methods. They measure the quality of the image explanation by comparing the target contours to the occluding contours of the current model instance θ, rendered onto the image domain with projection function

$$f_\theta \to C' = \{p'_1, p'_2, ..., p'_i, ...\}. \tag{5}$$

Our contour likelihood relies on soft correspondence, meaning for each point of the model proposed contour we search for the closest point in the jth target contour line

$$C_j = \{p^j_1, p^j_2, ..., p^j_i, ...\}. \tag{6}$$

To that end, the contours of the target and the current model instance are made invariant under $3D$ motion, so that only shape variation remains between the two. For this we find the rigid transformation between C' and C_j as:

$$h^j = \underset{h \in SE(3)}{\arg min} \|\overrightarrow{h(\{p^j_1, p^j_2, ..., p^j_i, ...\})} - \overrightarrow{\{f_\theta(p'_1), f_\theta(p'_2), ..., f_\theta(p'_i), ...\}}\|^2. \tag{7}$$

Our focus here is not on investigating the suitable likelihood function, thus we use the common assumption that the observed contour points are independent and are affected by additive Gaussian noise. Hence our likelihood rates the distances between closest points of the generated and the target contour after rigidly aligning them using a Gaussian noise model. We therefore define the likelihood as

$$l_j(\theta|C_j, \sigma_{c_j}) = \prod_i \mathcal{N}\left(f_\theta(p'_i) - h^j(p^j_i)|0, \sigma_{c_j}\right). \tag{8}$$

For N view registration, the final likelihood is computed as the product from N sets of contours:

$$l(\theta|C_1, ..., C_N, \sigma_{c_1}, ..., \sigma_{c_N}) = \prod_j^N l_j(\theta|C_j, \sigma_{c_j}). \tag{9}$$

2.4   Quantifying Uncertainty by Means of Sampling

In addition to measuring conventional mesh surface registration errors between the most likely reconstruction and the ground truth using both the average and the Hausdorff distance, the Bayesian sampling approach provides a means of measuring the uncertainty of the fit. We perform these measurements by considering all the samples that define our target distribution. We compute the

variance of each vertex point across these samples using the Frobenius norm, where we first compute the mean and the covariance for each direction in $3D$ coordinate space. Instead of having one most probable reconstruction, we infer the reduction in uncertainty in different regions of the bone and visualise it as a colour map.

To visualise our sample distribution and to investigate how the contour lines influences the reconstruction, we take slices of the samples across the length of the bone, the $z-$plane.

3   Experimental Setup and Results

A SSM of the femur is trained as is described in Sect. 2.1 using 113 bone surfaces. We use the process in Sect. 2.2 to develop synthetic contour data and thereafter show the applicability of the method on femur images. Our image data set consists of ten images simulated with a point rendering method through the perspective projection of ten $3D$ meshes. Half of the images are generated from meshes that were part of the model's training data, the other half is out of the model span.

To perform our experiment we use the Scalismo framework (Basel 2015) which is an open-source library for statistical shape modelling and model-based image analysis in the programming language Scala. This software already covers broad use cases for face image interpretation and 3D surface modelling. The framework is still under extension to allow X-ray simulations.

The $3D$ vertices of the model mean surface are passed through the contour projector described in Sect. 2.2 and projected to $2D$ for each X-ray viewing angle (see Fig. 1).
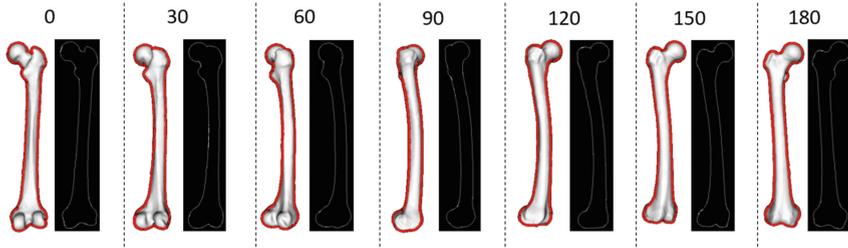

Fig. 1. Generation and projection of contour points from the 3D meshes.

The view-dependent contours are then fed as observations into the sampling algorithm described above. For a random proposal method, we consider 100 samples. From the resultant sample distribution (or surface meshes) we select one best posterior sample for each provided imaging angle. For the guided proposal method we use the best posterior of the random walk as the initial state of the model. We sample the step-size ϵ of the leapfrog step uniformly from

[0.000001, 0.0005). We also use a varying contour likelihood standard deviation σ to better represent different regions of the image. As such, areas with greater variance in the model have a higher σ. The number of steps denoted by L follows the Poisson distribution with Poisson(50). We use $n = 700$ HMC samples. The full experimental pipeline can be seen in the Fig. 2 below.
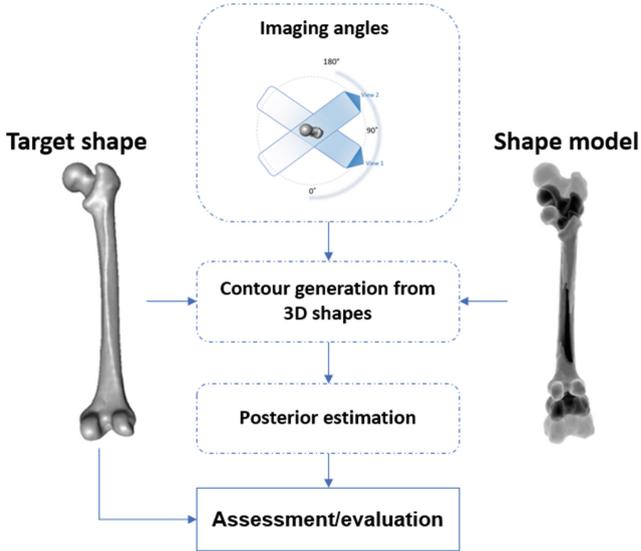

Fig. 2. Experimental setup.

We performed 2 experiments to evaluate different aspects of the proposed method. The first one evaluated the efficiency and accuracy of the proposed sampling algorithm with a single X-ray and the second one examined the feasibility of applying the method to a medical problem, that is $2D/3D$ reconstruction of bone surfaces with limited number of X-ray views [max n = 2].

3.1   2D/3D Reconstruction with Synthetic Data

Provided with an X-ray image contour projection, our task was to reconstruct a personalised $3D$ bone surface. Since the aim is to reduce the uncertainty in reconstruction using the limited number of X-ray images, it was important to assess the impact of the choice of the imaging angle in a setup with a single X-ray image and a pair of X-ray images (orthogonal to each other) on the quality and accuracy of the reconstructed surfaces.

Single X-Ray Registration. The single image experiment was performed to determine the optimal viewing direction for imaging a patient using a single X-ray source. The experiment was done for 7 different viewing directions with an

angle α ranging from 0 and $180°$ in an increment of 30. For each α there was one corresponding X-ray image to which we then registered a model and obtained the posterior distribution. The most probable shape from the distribution was used to represent the inferred $3D$ bone surface. The error between the inferred $3D$ shape and the ground truth was estimated with surface distance measures. The localised uncertainty reduction was estimated from the variance of the posterior sample distribution.

Reconstruction errors for all 7 angles are shown in Fig. 3 and in Fig. 4 - with an average across the samples. Both error metrics from the single X-ray image setting show that the worst angle for the $3D$ surface reconstruction was $90°$, which corresponds to a lateral view of the femur. This observation was confirmed in the variance diagram shown in the bottom left of Fig. 3. For each reconstructed surface, the low value was near 0 mm while the high value was near 5 mm. Therefore, the best imaging angles for the single X-ray source were 0 or $180°$ which is the anterior-posterior position of the femur bone. Even though these seemed to be the best angles, there were still regions of the bone with high variance.

Bi-Planar X-Ray Image Registration. Since the single X-ray image experiment did not significantly reduce the uncertainty in single image observations for our reconstruction, we investigated whether adding a second image to the reconstruction narrows the posterior distribution and/or reduces the variance. As for the previous experiment, we examined the uncertainty reduction in $3D$ reconstructions with an assumption that the second image captures information about the depth of the bone.

For this setup, an angle ϵ of separation between the images of the same pair was fixed to $90°$. The angle α between the pairs of images ranged from 0 to $180°$ in increments of 30 in order to find the combination that leads to the optimal reconstruction, that is, a reconstruction closest to the ground truth and having the least variance. The top right of Fig. 3 shows that the average distance error decreased with the addition of a second image for all the imaging angles. This is consistent with our expectation as the second image captures the depth of the bone and reduces variance. This was also evident in the reconstructed femurs on the bottom right of Fig. 3 where the shaft has less variance. However, the femoral head and the condyles were still regions of high uncertainty, hence the Hausdorff distance remained high. For these regions of the bone contour extraction becomes more complex because of self occlusion at the ends of the femur. The greater trochanter is superimposed with some part of the head and the intertrochanter region. When the contour extraction is applied it is only the outer contours that are detected rather than the contours from the regions of great interest. In the $3D$ view these contour lines are not connected or continuous.

4   Conclusion

We have presented a novel approach to $3D$ shape reconstruction from a pair of X-ray image contour projections, with application to the femur bone. While tra-
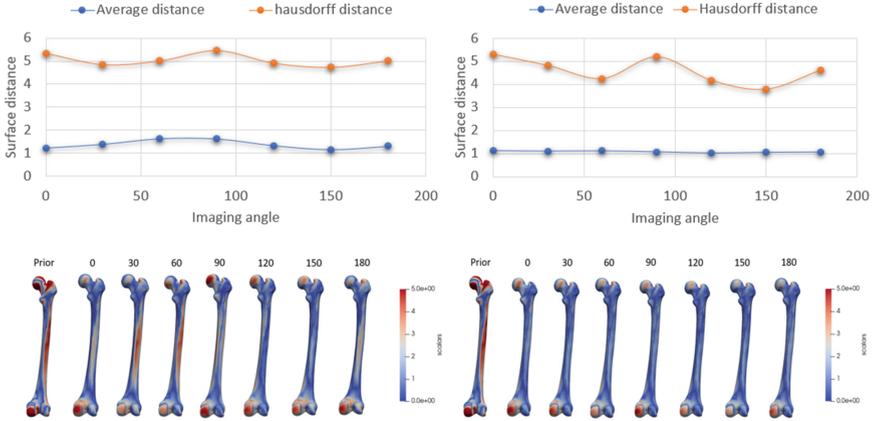
Fig. 3. Reconstruction results: the top row shows the surface distance between the reconstructed bone and the ground truth for a single X-ray experiment (left) and a bi-planar X-ray experiment (right), bottom row shows $3D$ reconstructed femurs for each viewing direction from a single X-ray experiment (left) and a bi-planar X-ray experiment (right). The colours represent the variance, red for high values and blue for low values. (Color figure online)
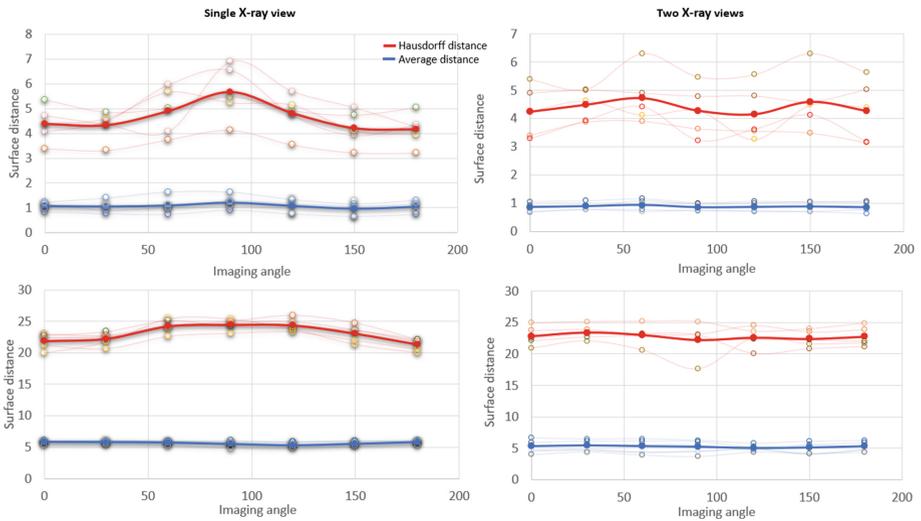

Fig. 4. Reconstruction results for a dataset with 10 femur bones: the top row shows the surface distance error between the reconstructed bone and the ground truth for the target images that are within the reach of the model, bottom row shows the surface distance error between the reconstructed bone and the ground truth for the target images that are out of the model reach. The bold colours shows an average across the samples while the light lines show each sample. (Color figure online)

ditional methods provide a single point-estimate, our approach uses Monte Carlo sampling techniques to recover the posterior distribution of plausible reconstructions. From the distribution, we were able to compute the amount of uncertainty that remains in the inferred $3D$ shape. Synthetic contour images and their corresponding ground-truth meshes were used to assess the feasibility of the method.

Our experiments indicate that using a single X-ray contour projection as an observation is insufficient for an accurate $3D$ surface estimate but, rather, that the $0°$ angle (anterior-posterior view) is most likely to provide a favourable first initialisation step of the model. Surface reconstruction accuracy improved the most when an angle-of-separation of $90°$ between the two X-ray views of 0 and $90°$ was used; the average error improved from $1.07\,\mathrm{mm}$ to $0.84\,\mathrm{mm}$. Regarding variance reduction, our experiments show a reduced uncertainty only along the shaft of the bone. This observation implies that adding a second view, with a $90°$ separation, does not significantly reduce the variance of highly curved regions. More experiments with different separation angles and different bones are necessary to reach a stronger conclusion regarding the suitability of our proposed method to 3D-2D reconstruction. Furthermore, since MC-based sampling allows different sources of information to be accommodated in its processing chain, our method can readily extend to intensity-based inferencing. We thus plan to investigate whether inner contours may help further constrain the model.

Once better established, we believe our proposed approach may benefit clinicians in two novel ways. First, the clinician is able to know the level of uncertainty with which the algorithm has estimated a bone surface. In addition, this type of understanding is possible for any local bone region; especially useful for presurgical planning which often focuses on a specific area. Finally, such insights can inform which viewing angles provide the least ambiguity, and are likely to improve the accuracy of bone surface reconstructions.

References

Baka, N., et al.: 2D–3D shape reconstruction of the distal femur from stereo X-ray imaging using statistical shape models. Med. Image Anal. **15**(6), 840850 (2011)

University of Basel. Scalismo—a scalable image analysis and shape modelling software framework available as open source (2015). https://github.com/unibas-gravis/scalismo

Duane, S., et al.: Hybrid Monte Carlo. Phys. Lett. B **195**(2), 216–222 (1987)

Egger, B.: Semantic Morphable Models. Ph.D. thesis. University of Basel (2017)

Fleute, M., Lavallée, S.: Nonrigid 3-D/2-D registration of images using statistical models. In: Taylor, C., Colchester, A. (eds.) MICCAI 1999. LNCS, vol. 1679, pp. 138–147. Springer, Heidelberg (1999). https://doi.org/10.1007/10704282_15

Hotelling, H.: Analysis of a complex of statistical variables into principal components. J. Educ. Psychol. **24**(6), 417 (1933)

Hurvitz, A., Joskowicz, L.: Registration of a CT-like atlas to uoroscopic X-ray images using intensity correspondences. Int. J. Comput. Assist. Radiol. Surg. **3**(6), 493 (2008)

Lüthi, M., et al.: Gaussian process morphable models. IEEE Trans. Pattern Anal. Mach. Intell. (2017)

Neal, R.M., et al.: MCMC using Hamiltonian dynamics. Handbook of Markov Chain Monte Carlo **2**(11), 2 (2011)

Sadowsky, O., Chintalapani, G., Taylor, R.H.: Deformable 2D-3D registration of the Pelvis with a limited field of view, using shape statistics. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) MICCAI 2007. LNCS, vol. 4792, pp. 519–526. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-75759-7_63

Schönborn, S., et al.: Markov chain Monte Carlo for automated face image analysis. Int. J. Comput. Vision **123**(2), 160–183 (2017)

Suter, T., et al.: Viewing perspective malrotation inuences angular measurements on lateral radiographs of the scapula. J. Shoulder Elbow Surg. **29**(5), 1030–1039 (2020)

Tu, Z., et al.: Image parsing: unifying segmentation, detection, and recognition. Int. J. Comput. Vision **63**(2), 113–140 (2005)

Zhu, S.C., Zhang, R., Tu, Z.: Integrating bottom-up/top-down for object recognition by data driven Markov chain Monte Carlo. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 738–745. IEEE (2000)

Zoppo, G.: Hamiltonian Monte Carlo for fitting point distribution models. Ph.D. thesis. University of Torino (2018)