Construction and validation of image-based statistical shape and intensity models of bone*

Cornelius Reyneke
Division of Biomedical Engineering
University of Cape Town
Cape Town, South Africa
RYNCOR001@myuct.ac.za

Xolisile Thusini
Division of Biomedical Engineering
University of Cape Town
Cape Town, South Africa
THSXOL001@myuct.ac.za

Tania Douglas
Division of Biomedical Engineering
University of Cape Town
Cape Town, South Africa
tania.douglas@uct.ac.za

Thomas Vetter
Department of Mathematics and Computer Science
University of Basel
Basel, Switzerland
thomas.vetter@unibas.ch

Tinashe Mutsvangwa
Division of Biomedical Engineering
University of Cape Town
Cape Town, South Africa
tinashe.mutsvangwa@uct.ac.za

***Abstract*—Three-dimensional models of bone structures, inferred from 3D and 2D imaging modalities, provide a number of uses for medical professionals. Such models are typically constructed using mesh-based approaches and some form of principal component analysis (PCA). Image-based approaches are understood to have a number of advantages over mesh-based ones, such as the absence of intrinsic segmentation as well as a better reproduction fidelity. Furthermore, Gaussian process morphable models have recently been shown to offer added benefits to the traditional PCA-based approach, such as the ability to easily incorporate prior knowledge into the model, and a resolution that can be made application-specific. We demonstrate how to build an image-based statistical shape and intensity model (SSIM) using Gaussian process morphable models and validate the quality of the model using common mesh model metrics that we adapt to the image-based paradigm. Our results show that the model can generate valid novel femur examples and can be registered to unseen femur examples with average root mean square error and average mutual information score of 0.172 and 0.644, respectively. However, only twenty binary training examples are used, each limited to contain approximately 65000 voxels. Future work will aim at extending the image-based SSIM to include the full range of CT intensity values; larger CT volumes and improve on the model building time.**

***Index Terms*—statistical shape and intensity model, image-based 3D/3D registration, Gaussian process morphable models**

I. Introduction

Medical professionals make use of three-dimensional (3D) reconstructions of patient-specific bone structures to detect bone-related pathologies, plan surgery, design implants and perform post-operative evaluations [1], [2]. Deformable models have been employed to aid in the segmentation process necessary for obtaining the patient-specific bone models from 3D volumes such as computed tomography (CT). Deformable models encode prior knowledge and assumptions about the typical 3D appearance of an anatomical structure which either rely on the intuitions of medical experts (analytical models) or on the statistics learnt from a set of labeled examples (statistical models) [3]. The parameters of the deformable model can then be manipulated to match the information gathered from a patient's images, be they 3D imaging modalities such as CT or magnetic resonance images (MRI), or 2D imaging modalities such as X-ray or ultrasound images, using some registration strategy. For a detailed discussion about deformable models in the context of medical image registration the reader is referred to [4]. A well-known and popular type of deformable model is a statistical shape model (SSM), which is particularly well-suited to the variability present in rigid structures such as bone [2], [5]. Statistical shape and intensity models (SSIMs) are similar to SSMs but also include intensity information, which enables additional clinical benefits, such as the estimation of the potential risk factors for individuals with osteoporosis (prediction of potential bone fractures and detection of existing vertebral fractures) and aiding in the often problematic task of distinguishing bone deformities from bone fractures [6], [7]. Typically however, the construction of SSIMs requires longer computation time and larger computer memory requirements [8] as it is necessary to establish correspondence not only on the surface of the shape, but also within the volume [2], [8]. Statistical shape and intensity models can be broadly categorized into two types of representations; the first models the volume as a mesh where the CT has been segmented as a set of nodes and mesh-morphing techniques are used to establish correspondence [9], [10]; the second models the volume as an image using the original CT voxels and correspondence is established using image registration algorithms [8], [11]. The researchers in [8] found that mesh-based SSIMs are better at reproducing shape information while image-based SSIMs was better at reproducing intensity information. An additional advantage of an image-based representation is that it does not explicitly require segmentation of the bone-of-interest, as is the case with the mesh-based approach [9]. This means that an

*Research supported by the SARChI programme of the National Research Foundation (South Africa).

entire CT volume can be modelled, which may be useful when the presence of surrounding tissue and structures can be leveraged to improve the segmentation process. Recently, Gaussian process morphable models (GPMMs) have been introduced for building SSMs and SSIMs while providing several added benefits [12]. For example, they provide a framework for easily incorporating both statistical relationships and expert intuitions regarding these relationships into a single deformable model [12]. Additionally, once built, a single GPMM allows resampling at arbitrary resolutions to suit different applications. This is useful for allowing dense correspondence of model features to be established without demanding unnecessary additional computational memory. The construction of a mesh-based SSIM of a bone has been demonstrated by [13], [14], [15], and a mesh-based SSM using GPMMs by [12]. An image-based SSIM was developed in [8]. However, we believe that an image-based SSIM built using GPMMs has not yet been demonstrated. In this paper we present an image-based SSIM of a CT image containing a femur bone using GPMMs. Furthermore, we report on the quality of the model by assessing both the validity of randomly generated instances of the model and how accurately the model can be fit to unseen examples. In this proof of concept, we limit the CT volumes used to develop the model to a relatively small size ($\sim 65k$ voxels in total) and also to binary intensity values, with the aim of demonstrating a full intensity distribution in future efforts. A total of twenty CT images comprise the training dataset, each of which is a $25 \times 25 \times 105$ volume of integer Hounsfield units (HU) ranging between $\{-1024; 2976\}$. We imposed these constraints in order to accommodate the large amount of time it takes to build and test a single statistical model.

II. Methods

A. Building the SSIM using Scalismo

Our method closely follows the procedure described by [12], except that the fundamental feature of representation is a voxel instead of a mesh point. The steps that are followed are shown in Fig. 1. These include: alignment of the training examples (Step 1), building a free-form deformation model (FFDM) from a single image (Step 2), using the FFDM to establish correspondence between the single image and the rest of the training images (Step 3), and finally, building the SSIM (Step 4).

We first aligned all the images in the dataset by manually labeling six landmarks on each image and then aligning them to those of the first image using generalized Procrustes analysis. Next, we build a FFDM using the first training image as a reference. A number of random samples are taken across the voxels of the reference image using a Gaussian kernel (amplitude $= 30$, sigma $= 50$), which is then used to build the FFDM as a preliminary Gaussian process. A non-rigid image-based 3D-3D registration schema is used to establish correspondence by iteratively deforming the FFDM to fit each of the rest of the training images (this is essentially the same process which is finally used to fit the model to an unseen femur image). In an image-based registration, we use the similarity of voxel intensity values to guide the registration process instead of the distance between mesh-points (as is typical of a mesh-based registration). The similarity of the fixed image and moving image is quantified based on the output of a similarity measure function. The goal is then to iteratively adjust the model parameters until the models current image instance (moving image) best matches a given training example (fixed image), as informed by the similarity measure (see Fig. 2). In order to gain a better understanding of this process the final set of model parameters can be considered such that they correspond to a deformation vector-field (DVF), and that when this DVF is applied to the reference image (which was used to build the FFDM) it results in a new image that optimizes the similarity measure function. By having knowledge of how each voxel of the reference image deforms to the respective voxels of each of the other training images, correspondence is established. Finally, principal component analysis (PCA) can be performed on the resulting set of DVFs, which results in the final GPMM being a SSIM. The same process described in Fig. 2 is used to fit the SSIM to an unseen femur image in order to reconstruct a patient-specific femur.

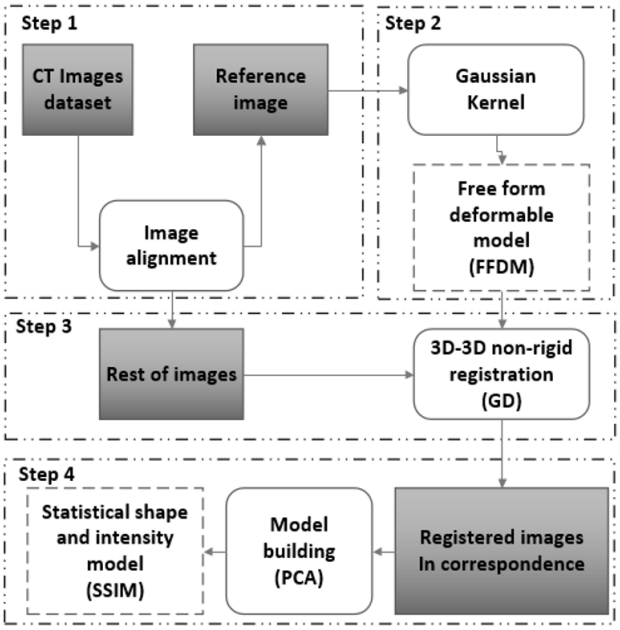

Fig. 1. Diagram illustrating the steps followed to build the SSIM from CT volumes. The dark shaded square-corned boxes are the persistent data, the dashed boxes are the models obtained, while the round-cornered boxes represent the algorithms that are executed. Step 1 is the preparation of the training dataset, Step 2 is the building of the FFDM, Step 3 is the 3D/3D non-rigid registration (for establishing correspondence), and Step 4 is the building of the SSIM, which encodes both shape and intensity information.

B. Experimental setup and validation protocol

The accuracy of the registration procedure depends strongly on the choice of optimization strategy and similarity measure. For our experiments we choose to use different registration configurations for establishing correspondence and for validating the model. This is in order to avoid the possibility

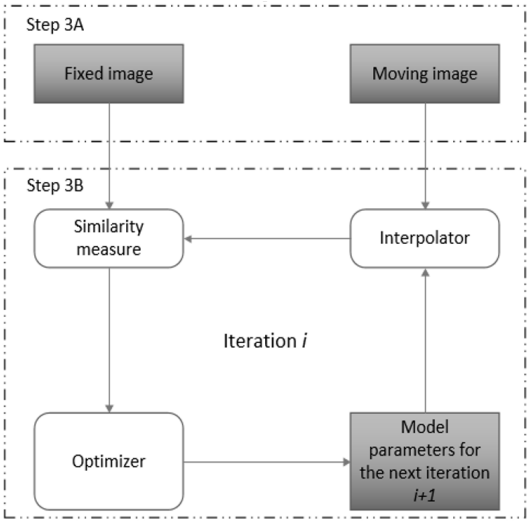

Fig. 2. An illustration of the algorithm used in the intensity-based registration methodologies. This is an elaboration of Step 3 of Fig. 1.

of biasing our results by merely over-fitting our model to the same measure used to gauge its quality. For establishing correspondence, we employ a stochastic gradient descent optimization strategy with a step-size of 0.05 and a total of 350 iterations per image, together with a mean squared error (MSE) metric. To improve computation time, the MSE is not calculated for each voxel of the volume, but rather on a randomly selected subset of them (sample size $= 1000$ voxels). The input images to the optimization strategy are also pre-processed; their range of values are normalized $\{0; 1\}$, subjected to a distance transform, as well as a Gaussian filter (sigma $= 1.5$) in order to avoid local minima when optimizing.

To validate the quality of the SSIM we make use of metrics typically computed for mesh-based shape models namely, specificity and generality [16], [17]. Specificity quantifies the ability of the model to generate instances similar to those available within the training set [16]. Generality is the ability of the model to generate instances not explicitly provided by the training set. This is an important characteristic of the model to measure because if the model is over-fitted to the training set, it may have excellent specificity but will hinder its ability to generalize to unseen examples [16]. In the mesh-based paradigm, specificity is computed as the average distance (such as root-mean-square error (RMSE) or mean-absolute-distance (MAD)) between randomly generated shapes and their nearest match in the training set, and generality is computed as the average distance between a certain number of unseen examples and their respective model reconstructions [16]. Typically, both specificity and generality are plotted as a function of the number of principal components employed by the model. For our purposes we make use of a similar protocol but in addition to RMSE we also compute mutual information (MI) values as an image-based counterpart to the Euclidean distance metrics used for mesh-model validation. Mutual information measures the amount of information that two data volumes in question contain about each other. The similarity is maximal when two volumes are perfectly aligned, whereas, when misaligned the MI will result in a value closer to zero [18]. The RMSE values that are reported are normalized such that they reflect grayscale intensity values instead of HU. Furthermore, instead of plotting the RMSE and MI values as a function of the number of principal components used, we plot these as a function of the number of images that the model is trained with (5, 10, 15, and 20). In doing so we are able to gauge to what extent additional training examples will improve the generality of the model. Once again, a stochastic gradient descent schema is used, with a step-size of 0.05 and 350 iterations. All the methods required for building a probabilistic morphable model from surface or volumetric meshes are provided by the Scalismo framework v0.16, an open-source software platform for shape modeling and fitting [19]. We have adapted these methods to accommodate CT volumes, which are comprised of voxels.

III. Results

Specificity and generality results are provided in Table I, and illustrated in Fig. 3 and Fig. 4. Since MI has not yet been utilized as a measure of specificity and generality, we included the mean MI (0.424 ± 0.12) and mean RMSE (0.247 ± 0.06) values among the images that comprise the training dataset, as a point of reference (larger MI values reflect a higher similarity, whereas the opposite is true for RMSE values).

TABLE I
Model Quality

Number of training images	RMSE		MI	
	Specificity	*Generality*	*Specificity*	*Generality*
5	0.158	0.232	0.583	0.470
10	0.177	0.190	0.563	0.602
15	0.180	0.172	0.581	0.637
20	0.177	0.172	0.583	0.645

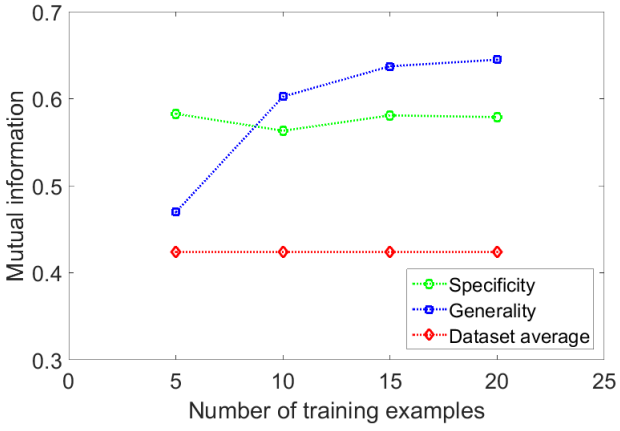

Fig. 3. Specificity and generality, measured using MI, as a function of the number of training examples used to build the SSIM.

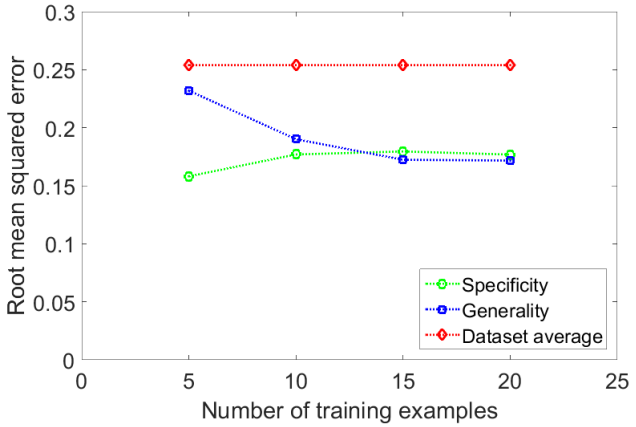

Fig. 4. Specificity and generality, measured using RMSE, as a function of the number of training examples used to build the SSIM.

By qualitatively evaluating the trend of MI values in Fig. 3 and comparing them to the RMSE values in Fig. 4, we believe it is reasonable to infer that the optimal number of training examples was between 10 and 15. This ensures a favorable balance between the specificity and generality, signaling that the model has not been over fitted. However, it does not necessarily follow that this guideline will still be valid when increasing the size of the CT volumes; this will be investigated in future work. The value for specificity falls within the range of MI values computed between the training examples, which suggests that randomly generated images are fairly typical when compared to the dataset, but slightly less diverse. As expected, the value for generality is better than for specificity, since the similarity measure was optimized in this case. While these values seem acceptable in this context, we are unable to come to robust conclusions concerning them since there are no other studies to compare them to. We plan to report registration accuracy in a similar fashion in future work, in order to objectively compare our results.

IV. Conclusion

We have demonstrated how to build an image-based SSIM of a femur bone CT image using GPMMs. We validated the quality of the model using an adaptation of conventional mesh model metrics to the image-based setting. The model was fitted to unseen femur examples with an average root mean squared error of 0.172, and a mutual information score of 0.644, and confirmed that randomly generated instances of the model maintain their validity as femur shapes.

The results obtained from our experiments show that image-based shape and intensity models of CT are a plausible alternative to the conventional mesh-based approach. Future efforts will extend the image-based SSIM to include the full range of CT intensity values, larger CT volumes, and to be used as the model component of a 2D/3D reconstruction algorithm.

Acknowledgment

The authors would like to thank Dr. M. Lüthi of the University of Basel for his valuable assistance and for providing the training dataset that was used in this study.

References

[1] G. Chintalapani, L. M. Ellingsen, O. Sadowsky, J. L. Prince, and R. H. Taylor, "Statistical atlases of bone anatomy: construction, iterative improvement and validation," *Med. Image Comput. Comput. Assist. Interv.*, vol. 10, no. Pt 1, pp. 499–506, 2007.

[2] N. Sarkalkan, S. Nazli, W. Harrie, and A. A. Zadpoor, "Statistical shape and appearance models of bones," *Bone*, vol. 60, pp. 129–140, 2014.

[3] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image-guided interventions," *Med. Image Anal.*, vol. 16, no. 3, pp. 642–661, Apr. 2012.

[4] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE transactions on medical imaging*, vol. 32, no. 7, pp. 1153–1190, 2013.

[5] G. Zheng, "Statistical shape model-based reconstruction of a scaled, patient-specific surface model of the pelvis from a single standard AP X–ray radiograph," *Med. Phys.*, vol. 37, no. 4, pp. 1424–1439, Apr. 2010.

[6] M. Roberts, E. Pacheco, R. Mohankumar, T. Cootes, and J. Adams, "Detection of vertebral fractures in dxa vfa images using statistical models of appearance and a semi-automatic segmentation," *Osteoporosis international*, vol. 21, no. 12, pp. 2037–2046, 2010.

[7] J. S. Gregory and R. M. Aspden, "Femoral geometry as a risk factor for osteoporotic hip fracture in men and women," *Medical engineering & physics*, vol. 30, no. 10, pp. 1275–1286, 2008.

[8] S. Bonaretti, C. Seiler, C. Boichon, M. Reyes, and P. Büchler, "Image-based vs. mesh-based statistical appearance models of the human femur: implications for finite element simulations," *Medical engineering & physics*, vol. 36, no. 12, pp. 1626–1635, 2014.

[9] L. Grassi, N. Hraiech, E. Schileo, M. Ansaloni, M. Rochette, and M. Viceconti, "Evaluation of the generality and accuracy of a new mesh morphing procedure for the human femur," *Medical engineering & physics*, vol. 33, no. 1, pp. 112–120, 2011.

[10] R. Bryan, P. S. Mohan, A. Hopkins, F. Galloway, M. Taylor, and P. B. Nair, "Statistical modelling of the whole human femur incorporating geometric and material properties," *Medical engineering & physics*, vol. 32, no. 1, pp. 57–65, 2010.

[11] F. P. Oliveira and J. M. R. Tavares, "Medical image registration: a review," *Computer methods in biomechanics and biomedical engineering*, vol. 17, no. 2, pp. 73–93, 2014.

[12] M. Lüthi, T. Gerig, C. Jud, and T. Vetter, "Gaussian process morphable models," *IEEE Trans. Pattern Anal. Mach. Intell.*, in press.

[13] G. Zheng, "Personalized X–ray reconstruction of the proximal femur via intensity-based non-rigid 2D–3D registration," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2011, pp. 598–606.

[14] J. Yao and R. H. Taylor, "Construction and simplification of bone density models," in *Medical Imaging 2001: Image Processing*, 2001.

[15] K. D. Fritscher, G. Agnes, and S. Rainer, "3D image segmentation using combined shape-intensity prior models," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 1, no. 6, pp. 341–350, 2007.

[16] M. A. Styner *et al.*, "Evaluation of 3D correspondence methods for model building," in *Biennial International Conference on Information Processing in Medical Imaging*. Springer, 2003, pp. 63–75.

[17] T. Mutsvangwa, V. Burdin, C. Schwartz, and C. Roux, "An automated statistical shape model developmental pipeline: application to the human scapula and humerus," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 4, pp. 1098–1107, 2015.

[18] L. Zollei, E. Grimson, A. Norbash, and W. Wells, "2D–3D rigid registration of X–ray fluoroscopy and ct images using mutual information and sparsely sampled histogram estimators," in *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, vol. 2. IEEE, 2001, pp. II–II.

[19] UNIBAS-GRAVIS, "Scalable image analysis and shape modelling," Mar 2018. [Online]. Available: https://github.com/unibas-gravis/scalismo/wiki/quickstart