2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Model Construction for the Estimation of Healthy Bone Shape and Density Distribution

DP Kramer, Johan Van der Merwe, and Marcel Lüthi

Abstract— **Statistical models are widely used within biomedical fields for automated segmentation and reconstruction of healthy geometry. In the absence of contralateral geometry, statistical models are a viable alternative for reconstructing healthy bone anatomy. Therefore, statistical models of shape and appearance were constructed from sample data based on the right femur of South African males, and their use in an automated segmentation and density estimation application was investigated. The models reproduced the shape and density distribution of the population with an average error of 1.3 mm and a 90% density fit. These results fall within the acceptable tolerance limits of reconstructive surgery and appear promising for practical use in implant design.**

Clinical Relevance— **Constructing and validating statistical models and registration algorithms provides the groundwork for further investigation into automating the digital reconstruction of pathological bone for use in implant design.**

I. Introduction

Segmental bone loss may result from high-energy trauma, tumor resection, bacterial infection, or revision surgery [1], [2]. Effective treatment requires access to specialist health care for surgical intervention [1], [3]. Delays in treatment could result in further bone loss, increased morbidity, and possible amputation [3].

Bone grafting (autografts) is the gold standard for treating and correcting segmental bone loss. This method is also used to evaluate the feasibility of other segmental bone loss treatments. However, autografting introduces a second surgical site for extracting bone tissue. This can result in additional complications such as extended recovery time, increased pain, and infection, thus increasing the patient's morbidity [4].

With the advancement of additive manufacturing, patient-specific implants (PSI) have become feasible and offer a modern alternative to autografts [2], [5]. When using PSIs for segmental bone loss treatment, potential advantages include reduced surgical time, improved patient outcomes, lower long-term costs, reduced recovery time, and no second morbidity [2]. The design of PSIs often requires knowledge of the healthy bone geometry prior to pathology. This information is conventionally inferred from specialist knowledge or, if available, a healthy contralateral computed tomography (CT) image [6], [7]. However, a healthy contralateral shape is not always readily available, and there could be significant differences in the bilateral anatomy of limb bones [8], [9].

Statistical shape models (SSM) could be a viable alternative to the use of contralateral bone geometry [9]. Statistical models are capable of representing the normal variability of an anatomical shape and density distribution within a specific population. Combining these models with automated segmentation algorithms and partial data from patients' pathological bone structures could produce a statistical shape and appearance estimate specific to the pathological bone presented [10], [11]. Statistical shape models are widely used to reconstruct bone shape, and are well studied in literature [11], [10]. However, statistical appearance models (AM), in which the variation between intensities of volumetric images are modeled, are computationally expensive to store and process, and fitting procedures are harder to implement successfully [11], [12]. This makes them inconvenient for use in practical design applications.

Therefore, in this paper we propose the use of an active shape model (ASM) to warp a single mean AM to a specific patient's geometry as a compromise between modeling accuracy and practicality. An SSM is constructed in order to develop an ASM with which to register segmented shapes onto unseen patient data. The deformations from the ASM provides a way to warp an image of a representative density distribution to the individual patient's anatomy that is sufficiently accurate for the purposes of surgical reconstruction.

II. Methods

A. Research Collection

The dataset for this study consisted of 50 retrospectively collected lower limb CT scans of South African males with an average age of 27.2 (±6.5). Samples were selected using the following inclusion criteria: 1) Patients must be between 18 and 45 years old; 2) have a healthy right femur; 3) have no radiologically identifiable prior disability; 4) obesity; 5) arthritis; or 6) other musculoskeletal conditions that may affect bone mineral density. CT scans were selected to have slice thicknesses and pixel sizes smaller than 1 mm. All scans were segmented by an industry expert. The sample size equals that of similar studies reported in literature [10], [13].

D. Kramer and J. v.d. Merwe are with the Department of Mechanical and Mechatronic Engineering, Stellenbosch University, Western-Cape, South-Africa (e-mail: 19865481@sun.ac.za, jovdmerwe@sun.ac.za).

M. Lüthi is with the Graphics and Vision Research Group, University of Basel, 4001 Basel, Switzerland (e-mail: marcel.luethi@unibas.ch)

The financial assistance of the South-African National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF.

B. Statistical Shape Model Construction

To represent the anatomical shape distribution within the population an SSM was trained through the Gaussian Process (GP) Morphable Models construction methodology proposed by Lüthi et al. [14]. The training data consisted of the 3D segmentations derived from the dataset (n = 50). First, a reference shape, $\Gamma_R = \{\boldsymbol{x} | \boldsymbol{x} \in \mathbb{R}^d\}$ with $\boldsymbol{x} = (x, y, z)$, representing the average shape of the sample data was assigned by computing the average size of the femur samples and selecting the closest sample. Secondly, using manually placed landmarks, Procrustes Analysis, and parametric registration procedures (see Section 4 of [14]), all remaining target shapes, $\Gamma_{1\to(n-1)} = \{\boldsymbol{x}_{1\to(n-1)} | \boldsymbol{x}_{1\to(n-1)} \in \mathbb{R}^d\}$, were aligned with the reference shape and correspondence was established. Thirdly, the deformations, $\boldsymbol{u} = \{\boldsymbol{u_1}, \ldots, \boldsymbol{u}_{n-1}\}$, of the target shapes from the reference shape were determined and the shape variability was established through the estimation of the mean deformation, $\boldsymbol{\mu}$, and covariance matrix, $\boldsymbol{k}$:

$$\boldsymbol{\mu}(\boldsymbol{x}) = \frac{1}{n-1}\sum_{i=1}^{n-1} \boldsymbol{u}_i(\boldsymbol{x}) \quad (1)$$

$$\boldsymbol{k}(\boldsymbol{x}, \boldsymbol{x}') = \frac{1}{n-2}\sum_{i=1}^{n-1} (\boldsymbol{u}_i - \boldsymbol{\mu}(\boldsymbol{x}))(\boldsymbol{u}_i - \boldsymbol{\mu}(\boldsymbol{x}))^T \quad (2)$$

Lastly, a GP model was defined, $\boldsymbol{u} \sim GP(\boldsymbol{\mu}, \boldsymbol{k})$, and a low-rank approximation of the covariance matrix was derived using Principal Component Analysis (PCA). The final shape model could then be described by:

$$\Gamma = \{\boldsymbol{x} + \boldsymbol{u}(\boldsymbol{x}) | \boldsymbol{x} \in \Gamma_R\} \quad (3)$$

with $\boldsymbol{u}(\boldsymbol{x})$ being a probabilistic entity for $\boldsymbol{u}: \Gamma_R \to \mathbb{R}^d$.

C. Active Shape Model Construction

An ASM combined with Bayesian model fitting using Markov Chain Monte Carlo (MCMC) was implemented for automated segmentation. Refer to Morel-Forster et al. [15] for a detailed explanation of the ASM and MCMC combination.

Following the procedures set out by Cootes et al. [11] an ASM for modeling local structures was trained in Scalismo, a scalable image analysis and shape modeling library [16]. Using the SSM and corresponding training samples established during registration, a set of uniformly distributed points, $\boldsymbol{x_{profiles}} = \{\boldsymbol{x_1}, \ldots, \boldsymbol{x}_i | x \in \Gamma_R\}$, on the surface boundary of the femur shape was randomly selected**.** For each of these points a line profile, $\boldsymbol{g_{profiles}} = \{\boldsymbol{g_1}, \ldots, \boldsymbol{g}_i\}$, of $2k + 1$ voxels was defined normal to the point surface, with k voxels on either side of the boundary surface. Using the corresponding training segmentations and CT images, the intensity profile, $\boldsymbol{g}_i$, was extracted for each selected point, $\boldsymbol{x}_i$, on each sample. A full multivariate GP, $\boldsymbol{g}_i \sim GP(\boldsymbol{\mu}_{gi}, \boldsymbol{k}_{gi})$, was then defined by applying the same principles used in (1) and (2) for each profile.

To obtain a good fit of the ASM, an appropriate initial starting point was required [11]. This was obtained by manually placing corresponding points on the ASM and unseen CT image. Rigid alignment and GP regression were then performed to fit the ASM to the CT image to get an initial starting point [14], [15]. MCMC combined with the Metropolis-Hastings algorithm was then used to propose potential translation, rotation, or shape corrections [15]. Using the same uniformly distributed points, $\boldsymbol{x_{profiles}}$, on the initial model instance, the intensity profiles, $\boldsymbol{g_{profiles}}$, of the unseen image were extracted and compared to the ASM profiles. Using prior knowledge of the shape and profile distributions captured within the ASM, the likelihood of the suggested translation, rotation, or shape correction was calculated (similar to [15]). If the correction was more likely than the current model instance, the shape and pose were corrected, another correction was suggested and the cycle repeated [15], [16].

D. Appearance Model Construction

For simplicity, only a mean intensity model was constructed to represent the anatomical density distribution within the population. Using interpolation, the deformations, $\boldsymbol{u}$, calculated for each target shape during registration, were extended to the entire CT image for each sample. This warped the voxel data of each CT image to the reference shape, Γ_R. With the sample images shape-normalized, the corresponding intensity vector, $\boldsymbol{g}_{im}$, from each femur image could be extracted [11]. Note that the voxel data contained in $\boldsymbol{g}_{im}$ for each image contained only the femur bone intensity data.

To combat intensity variation from different machine imaging settings the intensity data contained within $\boldsymbol{g}_{im}$ was normalized to the mean intensity vector, $\bar{\boldsymbol{g}}$ [10], [11]. However, with $\bar{\boldsymbol{g}}$ of the sample population not yet known at this stage, Cootes et al. [11] proposed the following process: 1) Assign one of the samples as the initial $\bar{\boldsymbol{g}}$. 2) Normalize the intensity values of all samples to $\bar{\boldsymbol{g}}$. 3) Calculate the new $\bar{\boldsymbol{g}}$ of the normalized samples. 4) Iterate until convergence.

In this study, the femur image of the reference shape was selected as the initial $\bar{\boldsymbol{g}}$. Vector zero mean unit standard deviation was selected to normalize the intensity values of all samples. The convergence criterion was the average difference between the elements of the old and new $\bar{\boldsymbol{g}}$ of each iteration. An average difference value of 0.1 yielded good results. After the final $\bar{\boldsymbol{g}}$ was calculated it was used to generate a mean CT image, Fig 1 (top).

E. Validation Procedures

For validation of the models constructed and segmentation algorithms used, leave-one-out testing was performed on the same randomly selected 10 samples throughout. Thus, 10 additional models were constructed for the SSM, ASM, and AM. In each successive model, the sample used for testing was omitted from the model's training data during construction.

When validating the SSM, the model generalization was tested, giving an account of how well the model could describe members of the statistical population outside of the sample data used to train the model [13]. The 10 SSMs for leave-one-out testing were fitted to the samples left out of the training set, with no pre-established correspondence. During the fitting process a parametric, non-rigid registration algorithm was used to fit the model to the unseen sample

shape [16]. To evaluate the fit, two metrics were calculated. The first, was the average distance error, where for each point on the unseen shape the shortest distance to the fitted model surface was calculated, and the overall average was returned. Secondly, the Hausdorff distance between the surface of the unseen shape and the fitted model was calculated.

TABLE I. RESULTS OF MODEL VALIDATON

Sample	SSM		ASM		Mean Density	
	Avg. Dist. [mm]	Hdf. Dist. [mm]	Avg. Dist. [mm]	Hdf. Dist. [mm]	R2 [-1, 1]	MAE [HU]
NS_401	0.56	3.28	1.27	5.99	0.93	118
PB_126	0.80	4.56	1.55	7.64	0.89	154
SD_235	0.55	3.39	1.08	4.60	0.91	135
SM_517	0.58	3.30	1.44	8.61	0.93	124
SN_234	0.65	4.02	1.12	6.07	0.91	139
SZ_234	0.75	4.61	1.00	6.03	0.87	169
TC_603	0.63	4.78	1.23	6.34	0.94	112
TG_375	0.58	3.54	1.14	6.56	0.91	137
TL_126	0.52	3.87	1.54	6.75	0.85	174
TS_710	0.71	4.37	1.79	7.96	0.85	177
Avg.	0.63	3.97	1.32	6.66	0.90	144

The 10 leave-one-out ASMs were fitted to their corresponding unseen CT image. Refer to algorithm 1 of [15] for the detailed fitting process. To evaluate the fit, the same two metrics were used as in the SSM validation. Thus, the automated ASM segmentation was compared to the manual segmentation done by an industry expert.

To validate the AM, the 10 unseen images were again shape normalized to the reference shape using the interpolated deformations $\boldsymbol{u}$, as discussed in the *Section B*. From the shape-normalized images the intensity vector $\boldsymbol{g}_{im}$, of each image was extracted and vector normalized to $\bar{\boldsymbol{g}}$. From the intensity vectors the mean absolute error (MEA) and the overall fit of the mean through the adjusted R square (R^2) metric was calculated. As a visual aid, the univariate standard deviation of the entire sample set was calculated and displayed using a generated CT image. This aids in understanding the location of the estimated density distribution error within the femur bone structure.

III. RESULTS

A. Statistical Shape Model Validation

Table 1 shows that the average distance error for all samples tested was below 1 mm, averaging at 0.63 mm. Through visual inspection, the Hausdorff distance for each testing sample was found to occur at the proximal or distal end of the femur. In most cases, this was at the trochanteric fossa due to the anatomical variance of the trough. Visual inspection also showed that the diaphyseal and metaphyseal areas of the femur were fitted more accurately than the epiphyseal areas.

B. Active Shape Model Validation

When validating the ASM the average distance error for all samples was below 2 mm, averaging at 1.32 mm. Again, the Hausdorf distance was inspected visually and observed to occur on the proximal or distal ends of the femur. This is most likely due to neighboring bone geometry or protrusions at the joint areas within the CT image. The ASM profiles at these areas were fitted to the local minima of the bony protrusions close to the joints and not the femur ends [15]. This could possibly be prevented through a better initial starting position, manual intervention during the fitting process, or an added MCMC parameter [15].

C. Appearance Model Validation

Figure 1 shows the mean image (top) and standard deviation (bottom), which ranges in Hounsfield Units (HU) from 0 to 2342 and 0 to 715 respectively. Most of the variation occurs within the medullary cavity as well as the proximal and distal ends.

When analyzing the R^2 metric in Table 1, it becomes apparent that the mean AM gave a good estimation of the intensity distribution for the test samples, averaging a 90% fit. For all the voxel data in $\bar{\boldsymbol{g}}$, the average MAE makes up 40% of the lower quartile (321 HU), 25% of the median quartile (545 HU), 10% of the upper quartile (1230 HU), and 5% of the upper cortical bone limit (2342 HU).

IV. DISCUSSION

Audenaert et al. [13] showed that SSMs constructed from 50 training femurs can have a model generalization error below 1 mm. It is clear from Table 1 that the full SSM should have a generalization error around 0.63 mm. This shows that the constructed SSM is able to describe members of the South African male population outside of the sample data used to train the model. Audenaert et al. [13] also found that the generalization error for their femur models did not go below 0.5 mm with a sample size of 250. Though the generalization error is also dependent on the statistical population, increasing the dataset would most likely have a limited effect on the generalization of the SSM.

When analyzing the results for the ASM it was found that the average distance error for each sample was between 1 and 2 mm, an acceptable error for reconstruction surgery as observed in literature [17], [18]. The segmentation was also done automatically without any specialist knowledge and within a few seconds.

In Fig. 1 (bottom) it can be observed how the mean AM represents the intensity distribution of the sample set. The largest standard deviations were found within the medullary cavity and epiphyseal ends, most likely due to the dynamic nature of cancellous bone tissue [19]. The relatively low standard deviation of the compact cortical bone, especially in the diaphyseal area is indicative of an accurate representation by the mean image.

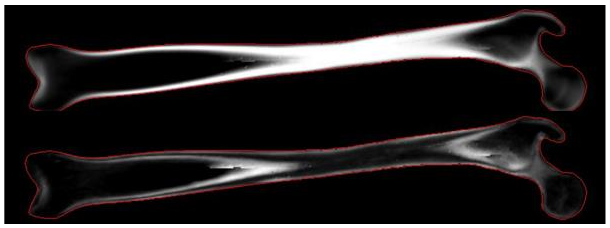

Figure 1. An image of the mean AM (Top) followed by its univariate standard deviation (Bottom).

In literature, AMs of bone have been used for two applications. Firstly, to assess risk factors for fractures in long bone structures and joint anatomy related to osteoporosis [12]. Secondly, to calculate the stress and strain distribution, study bone tissue adaptation and simulate cycle loading on implants via patient-specific finite element (FE) models [12], [20]. For both applications, an AM or FE model is populated by estimating the bone mineral density (BMD) of the patient from *in vivo* imaging using several direct methods [20], [21] [22]. Lekadir et al. [22] describe a novel method for estimating the trabecular micro-architecture from healthy CT images using a combined statistical shape and appearance model, trained from *ex vivo* micro-CT images. However, while direct methods are computationally less expensive they do not incorporate prior knowledge of variation within a population to guide the end result towards a statistically valid outcome. Secondly, while a statistical AM can accurately model the variation within a population, training it requires a healthy high-resolution training database and it is computationally much more expensive during data processing, model construction, and model fitting [11], [12], [22].

Thus, the mean AM used and validated during this study is a useful compromise. When estimating the density distribution, we get the speed of the direct methods with the additional statistical validity of a partial AM, as seen from the average 90% density fit throughout testing. Furthermore, the appearance estimate based on CT image intensity can be directly linked to BMD, which can be used to estimate the Young's modulus and ultimate strength of bone [21]. Thus, even with some error in the estimation, the density distribution captured within the mean AM could enable the design of patient-specific functionally graded lattice structures that closely resemble the mechanical properties of an individual's healthy bone structure to avoid stress shielding and stimulate new bone growth and nutrient flow at the implant site [2], [6], [23], [13].

V. Conclusion

This study presented a method in which a representative mean density image is warped to a patient's image with the use of an active shape model. The results are promising, with accuracies within acceptable tolerances for the purposes of an application aimed towards the surgical reconstruction of long bone defects. Future work would focus on the estimation and reconstruction of pathological geometry based on sparse inputs as is normally seen in patients suffering from long bone defects.

References

[1] W. Wang and . K. W. Yeung, "Bone grafts and biomaterials substitutes for bone defect repair: A review," *Bioactive Materials,* vol. 2, no. 4, pp. 224--247, 2017.

[2] X. Wang, S. Xu, S. Zhou, W. Xu, M. Leary, P. Choong, M. Qian, M. Brandt and Y. M. Xie, "Topological design and additive manufacturing of porous metals for bone scaffolds and orthopaedic implants: A review," *Biomaterials,* vol. 83, pp. 127-141, 2016.

[3] L. Marais, N. Ferreira, C. Aldous and T. Le Roux, "The pathophysiology of chronic osteomyelitis," *SA Orthopaedic Journal,* vol. 12, no. 4, pp. 14--18, 2013.

[4] C. Mauffrey, B. Barlow and W. Smith, "Management of segmental bone defects," *JAAOS-Journal of the American Academy of Orthopaedic Surgeons,* vol. 23, no. 3, pp. 143--153, 2015.

[5] A. A. Zadpoor and H. Weinans, "Patient specific bone modeling and analysis: the role of integration and automation in clinical adoption," *Journal of biomechanics,* vol. 48, no. 5, pp. 750--760, 2015.

[6] K. Tetsworth, S. Block and V. Glatt, "Putting 3D modelling and 3D printing into practice: virtual surgery and preoperative planning to reconstruct complex post-traumatic skeletal deformities and defects," *SICOT-J,* vol. 3, 2017.

[7] C. Letta, A. Schweizer and P. Furnstahl, "Quantification of contralateral differences of the scaphoid: a comparison of bone geometry in three dimensions," *Anatomy research international,* vol. 2014, 2014.

[8] B. M. Auerbach and C. B. Ruff, "Limb bone bilateral asymmetry: variability and commonality among modern humans," *Journal of human evolution,* vol. 50, no. 2, pp. 203--218, 2006.

[9] F. Mauler, C. Langguth, . A. Schweizer, L. Vlachopoulos, T. Gass, M. Luthi and P. Furnstahl, "Prediction of normal bone anatomy for the planning of corrective osteotomies of malunited forearm bones using a three-dimensional statistical shape model," *Journal of Orthopaedic Research,* vol. 35, no. 12, pp. 2630--2636, 2017.

[10] T. Heimann and H.-P. Meinzer, "Statistical shape models for 3D medical image segmentation: a review," *Medical image analysis,* vol. 13, no. 4, pp. 543--563, 2009.

[11] T. F. Cootes, C. J. Taylor and others, Statistical models of appearance for computer vision, Technical report, University of Mancheste, 2004.

[12] . N. Sarkalkan, H. Weinans and A. A. Zadpoor, "Statistical shape and appearance models of bones," *Bone,* vol. 60, pp. 129--140, 2014.

[13] E. A. Audenaert, C. Pattyn, G. Steenackers, J. De Roeck, D. Vandermeulen and P. Claes, "Statistical shape modeling of skeletal anatomy for sex discrimination: their training size, sexual dimorphism, and asymmetry," *Frontiers in bioengineering and biotechnology,* vol. 7, p. 302, 2019.

[14] M. Luthi, T. Gerig, C. Jud and T. Vetter, "Gaussian process morphable models," *IEEE transactions on pattern analysis and machine intelligence,* vol. 40, no. 8, pp. 1860--1873, 2017.

[15] A. Morel-Forster, T. Gerig, M. Luthi and T. Vetter, "Probabilistic fitting of active shape models," *International Workshop on Shape in Medical Imaging,* pp. 137--146, 2018.

[16] The Graphics and Vision Research Group at the University of Basel, "Scalismo," 2021. [Online]. Available: https://scalismo.org/. [Accessed 17 April 2021].

[17] K. Weise, F. Kock, C. Blake, T. Franken and J. Jordaan, "The accuracy of pre-operative digital templating in total hip arthroplasty performed in a low-volume, resource-constrained orthopaedic unit," *SA Orthopaedic Journal,* vol. 19, no. 1, pp. 28--32, 2020.

[18] R. H. Schepers, G. M. Raghoebar, A. Vissink, M. W. Stenekes, J. Kraeima, J. L. Roodenburg, H. Reintsema and M. J. Witjes, "Accuracy of fibula reconstruction using patient-specific CAD/CAM reconstruction plates and dental implants: a new modality for functional reconstruction of mandibular defects," *Journal of Cranio-Maxillofacial Surgery,* vol. 43, no. 5, pp. 649--657, 2015.

[19] J. G. Betts, Y. Kelly A., W. James A., J. Eddie, P. Brandon, K. Dean H., K. Oksana, J. Jody E., W. Mark and D. Peter, Anatomy and Physiology, 2020 PDF Edition ed., OpenStax, 2013.

[20] S. Poelert, E. Valstar, H. Weinans and A. A. Zadpoor, "Patient-specific finite element modeling of bones," *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine,* vol. 227, no. 4, pp. 464--478, 2013.

[21] L. Duchemin, V. Bousson, C. Raossanaly, C. Bergot, J. Laredo, W. Skalli and D. Mitton, "Prediction of mechanical properties of cortical bone by quantitative computed tomography," *Medical engineering \& physics,* vol. 30, no. 3, pp. 321--328, 2008.

[22] K. Lekadir, J. Hazrati-Marangalou, C. Hoogendoorn, Z. Taylor, B. van Rietbergen and A. F. Frangi, "Statistical estimation of femur micro-architecture using optimal shape and density predictors," *Journal of biomechanics,* vol. 48, no. 4, pp. 598--603, 2015.

[23] M. Long and H. Rack, "Titanium alloys in total joint replacement, a materials science perspective," *Biomaterials,* vol. 19, no. 18, pp. 1621--1639, 1998.