

On predicting 3D bone locations inside the human body

Abdelmouttaleb Dakri^{1*}, Vaibhav Arora^{1*}, Léo Challier¹, Marilyn Keller²,
Michael J. Black², and Sergi Pujades¹

¹ Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, France

² Max Planck Institute for Intelligent Systems, Germany

Abstract. Knowing the precise location of the bones inside the human body is key in several medical tasks, such as patient placement inside an imaging device or surgical navigation inside a patient. Our goal is to predict the bone locations using only an external 3D body surface observation. Existing approaches either validate their predictions on 2D data (X-rays) or with pseudo-ground truth computed from motion capture using biomechanical models. Thus, methods either suffer from a 3D-2D projection ambiguity or directly lack validation on clinical imaging data. In this work, we start with a dataset of segmented skin and long bones obtained from 3D full body MRI images that we refine into individual bone segmentations. To learn the skin to bones correlations, one needs to register the paired data. Few anatomical models allow to register a skeleton and the skin simultaneously. One such method, SKEL, has a skin and skeleton that is jointly rigged with the same pose parameters. However, it lacks the flexibility to adjust the bone locations inside its skin. To address this, we extend SKEL into SKEL-J to allow its bones to fit the segmented bones while its skin fits the segmented skin. These precise fits allow us to train SKEL-J to more accurately infer the anatomical joint locations from the skin surface. Our qualitative and quantitative results show how our bone location predictions are more accurate than all existing approaches. To foster future research, we make available for research purposes the individual bone segmentations, the fitted SKEL-J models as well as the new inference methods at <https://3dbones.is.tue.mpg.de>.

1 Introduction

Knowing the precise location of the bones inside the human body is key in several medical tasks, such as patient placement inside an imaging device, surgical navigation inside a patient and in biomechanics to compute the joint kinematics accurately. Recent CT scanning devices (e.g. Siemens Healthineers and GE Healthcare) are equipped with RGB or depth cameras to enhance patient positioning [1–3]. Accurate patient placement can reduce or eliminate ionizing scout images, thus reducing the overall radiation dose of the patient.

* Equal contribution

The key is to be able to predict the location of internal structures from external surface observations. The research community has addressed this in several ways. Several graphics approaches [4–9] deform an anatomic atlas to match the skin observation, but these methods are not validated against clinical data and may create incorrect anatomies. Anatomic models learned from cohorts of clinical data have been recently proposed [10, 11]. BOSS [10] is a statistical skin-bone-organs model learned from 3D medical data. Fitting the model to match novel data requires expensive optimization, which is complicated by the fact that the skin and skeleton use different kinematic trees. Anatomy Completor [11] can complete the shape of missing organs from the shape of the neighboring ones but does not predict them from a surface observation.

Several approaches have specifically targeted the task of inferring the internal anatomy from surface measurements. For example, Teixeira et al. [12] generate 2D X-rays from surface observations and Wu et al. [13] go further in that they generate a volumetric CT phantom. In A-SMPL [14], the authors predict the locations of the lungs as the patient moves, but the initial model is created from an actual scan. In OSSO [15], the authors learn how to predict the 3D bone locations inside the body from 2D DXA images [16]. SKEL [17] uses a different strategy to learn the skin-skeleton relationship, by leveraging vast amounts of motion capture data [18] and biomechanical physics simulation [19, 20]. The recent HIT approach [21] predicts the location of several tissues (lean, adipose, and bone) inside the body from a surface observation, but their bone prediction is binary and does not include individual bone segmentations. In our work, we enrich the HIT dataset with multi-bone labels and learn how to predict them from the body surface. All existing approaches predicting the bone locations from a surface observation [12, 15, 17, 21] validate their predictions: i) on 2D data (X-rays, DXA) [12, 15] suffering from a 3D-2D projection ambiguity; or ii) with surface motion sequences and biomechanical methods [17], but this only provides pseudo ground truth. Only HIT [21] and Wu et al. [13] validate their predictions on 3D imaging data.

In this work, we build on the HIT dataset [21], by using their 3D segmentations of skin and bones, obtained from 3D full-body MRI images. We refine their binary bone segmentation data into individual bone segments, namely the humerus, radius-ulna, pelvis, femur, and tibia-fibula. These individual segmentations are key in two aspects. First, their fine grained semantics allow us to go beyond the binary bone prediction of HIT and predict anatomically precise structures to create an articulated digital twin. Secondly, they provide the means to have a deeper insight on the skeleton predictions, allowing to quantify which methods perform best in which body part.

To effectively obtain more accurate predictions, we register the biomechanical model SKEL [17] to the individually segmented bones and skin. However, this is not straightforward: given a skin shape, the skeleton estimated by SKEL is deterministic: only one skeleton is possible, whereas the same skin could embody different bone shapes underneath [22]. We thus extend SKEL into SKEL-J, by adding an over-parametrization of the joint locations, to accurately register

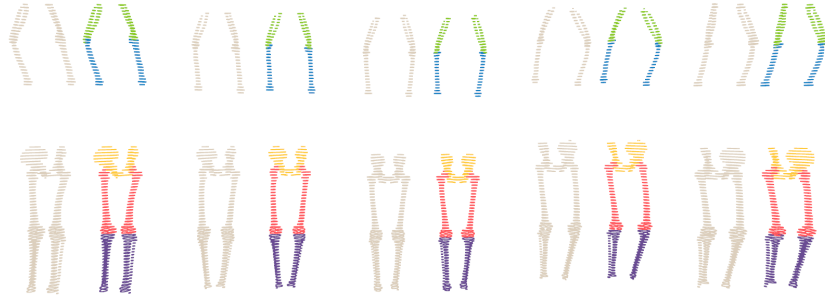


Fig. 1: Five subjects pairs of: (left) HIT [21] bone segmentations, (right) our multi-bone segmentations: radius-ulna, humerus, pelvis, femur and tibia-fibula.

the actual skin and bones data. We effectively adjust the joint locations defining the kinematic tree as well as the pose parameters to match the segmented bones. With these registrations, we can learn a more accurate model relating the human body surface and the underlying bones. Our refined SKEL-J model has an improved joint regressor learned from and validated on segmented medical scans. Our quantitative evaluation shows that our improved model predictions are more accurate than existing state-of-the-art approaches.

In summary, we contribute i) a dataset of segmented individual bones paired with skin observations; ii) SKEL-J: an improved version of SKEL and a method to register it simultaneously to multiple bones and skin; iii) the corresponding registrations matching the segmented structures; and iv) new regressors that more accurately infer the skeleton from a skin observation. The individual bone segmentations, the improved SKEL-J fits, as well as the new inference methods are made available for research purposes at <https://3dbones.is.tue.mpg.de>.

2 Method

We first introduce the dataset we use and how we refine the bone segmentations into individual bones. Then we detail our *multi-bone and skin registration* as well as how we learn the relation between the skin and the bone locations.

Dataset. We start with the HIT dataset [21] consisting of 381 white German subjects (235 females, 146 males). Their age range spans 23 to 65 years with 4 balanced groups (A: < 35 y; B: 35–44 y; C: 45–54 y, D: > 54 y;) and three BMI groups (normal weight (≤ 25), overweight (25–29.9), and obese (≥ 30)). All subjects were considered healthy according to physical examination and routine laboratory tests. For each subject i , a full body MRI segmentation is provided, as well as the skin point-cloud $\mathcal{S}_i$ and the SMPL model mesh fits $\mathbf{S}_i$ matching the skin $\mathcal{S}_i$ of the subject. The pixel size of the original MRI images is 2.0×2.0 mm, and the slice thickness is 10 mm.

3D Individual Bone extraction. In the HIT dataset [21], the bones have a binary label (bone/not bone). We further manually annotate five individual bones (humerus, radius-ulna, pelvis, femur, and tibia-fibula) for 10 cases. From these annotations, we train and refine a nnUnet [23] with an iterative process with the help of human supervision [21, 24, 25] and run it on the full dataset. From the segmented images, we extract the bone contours and generate 3D point-clouds for each individual bone. We note $\{\mathcal{B}_i^*\}$ the set of five bone point-clouds, where $\mathcal{B}_i^j$ with $j \in \{1, \dots, 5\}$ refers to the individual bone index and i identifies the subject. Fig. 1 illustrates the obtained multi-label bone point-clouds for 5 cases. We consider the point-clouds sets $\{\mathcal{B}_i^*\}$ to be ground truth.

2.1 3D bones and skin registration

While the bone segmentations $\{\mathcal{B}_i^*\}$ are accurate and allow us to identify their location inside the body, they only yield sparse and unstructured information. Directly predicting a 3D bone point cloud that accurately matches $\{\mathcal{B}_i^*\}$ from the body surface is highly challenging. To make the problem more tractable, we focus on predicting a more straightforward parametric model of the bones, like OSSO [15] and SKEL [17] do. To that end, for each subject i , we want to compute SKEL parameters, so that the skin and bones of SKEL accurately match the segmented point-clouds $\mathcal{S}_i$ and $\{\mathcal{B}_i^*\}$. The result of this registration provides us with paired data to learn their relationship.

The SKEL model [17] has shape $\beta \in \mathbb{R}^{10}$ and pose $\mathbf{q} \in \mathbb{R}^{46}$ parameters which define both a skeleton and a skin mesh. We note the 3D vertices of those meshes respectively $\mathbf{v}_{\text{skel}}(\beta, \mathbf{q}) \in \mathbb{R}^{3 \times 17890}$ and $\mathbf{v}_{\text{skin}}(\beta, \mathbf{q}) \in \mathbb{R}^{3 \times 6890}$. The SKEL skin equation follows Eq. 5 and 6 from the SMPL model paper [26], deforming a template mesh vertices $\mathbf{T}$ with learned shape basis coefficients $\mathbf{B}$ controlled by the β parameter. From these shaped vertices, the SKEL joints $\mathbf{J}^{\text{SKEL}} \in \mathbb{R}^{3 \times 24}$ are computed using the regressor $\mathcal{J}_{\text{SKEL}}$ learned from mocap data and biomechanical fits. To simplify the notation, we write $\mathbf{J}^{\text{SKEL}} = \mathcal{J}_{\text{SKEL}}(\beta)$, as $\mathcal{J}$ operates on vertices which depend on β . Given a pose $\mathbf{q}$, pose dependent deformations $\mathbf{B}_P(\mathbf{q})$ are added to the shaped vertices, and linear blend skinning is used to obtain the posed vertices. The skeleton vertices undergo a similar process, with the difference that no pose dependent deformations are added. Additionally, we note the posed joint locations, posed with the parameter $\mathbf{q}$, as $\mathbf{J}_\mathbf{q}$.

Improving SKEL. As mentioned before, the SKEL model was not validated against actual 3D clinical data, thus the link between the skin and the skeleton has limited accuracy, i.e. the locations of the bones inside the skin could not accurately match the actual data. Moreover, one limitation of SKEL is that given a skin shape, the estimated internal skeleton is deterministic: only one skeleton is possible. However, the same skin could embody different bone shapes underneath [22]. To improve SKEL, we introduce an additional degree of freedom: $\Delta \mathbf{J} \in \mathbb{R}^{3 \times 46}$ which allows to add an offset to the regressed joint locations $\mathbf{J}^{\text{SKEL}}$. The effect of this offset is two-fold: it effectively modifies the kinematic tree used for the linear blend skinning, and also affects the size of the bones related to this joint, e.g. if the knee joint is moved up, the tibia and fibula become

longer, whereas the femur becomes shorter. This modification does not affect the shape of the skin but allows to adjust the shape of the bones and the joint locations inside the body. We name this improved model SKEL-J and note its skeleton and skin vertices as $\mathbf{v}_{\text{skel}}(\boldsymbol{\beta}, \mathbf{q}, \Delta\mathbf{J})$ and $\mathbf{v}_{\text{skin}}(\boldsymbol{\beta}, \mathbf{q}, \Delta\mathbf{J})$ respectively.

Registration. With SKEL-J we can compute the parameters $(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)$ for each subject, so that $\mathbf{v}_{\text{skel}}(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)$ matches $\{\mathcal{B}_i^*\}$ and $\mathbf{v}_{\text{skin}}(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)$ matches $\mathbf{S}_i$. Note that we use $\mathbf{S}_i$ as a surrogate of the subjects' skin S_i . We define the bone loss

$$L_{\text{bone}}(\{\mathcal{B}_i^*\}; \boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i) = \sum_{j \in \{1, \dots, 5\}} p2m(\mathcal{B}_i^j, \mathbf{v}_{\text{skel}}^j(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)), \quad (1)$$

where $p2m(., .)$ is the point-to-mesh distance for all bone scan vertices, and $\mathbf{v}_{\text{skel}}^j$ refers to the skeleton vertices associated with bone j . We define the skin loss as

$$L_{\text{skin}}(\mathbf{S}_i; \boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i) = p2m(\mathbf{S}_i, \mathbf{v}_{\text{skin}}(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)), \quad (2)$$

and the total loss

$$L(\mathbf{S}_i, \{\mathcal{B}_i^*\}; \boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i) = L_{\text{skin}}(\mathbf{S}_i; \boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i) + \lambda L_{\text{bone}}(\{\mathcal{B}_i^*\}; \boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i). \quad (3)$$

Variables after “;” in the loss input indicate that they are being optimized for, i.e. for each subject i , we compute the parameters $(\boldsymbol{\beta}_i, \mathbf{q}_i, \Delta\mathbf{J}_i)$ that minimize Eq. (3) and obtain posed skin and skeleton vertices $(\mathbf{v}_{\text{skin}}^i, \mathbf{v}_{\text{skel}}^i)$. We name the obtained unposed joint locations $\mathbf{J}_i^{\text{GT}} := \mathbf{J}_i^{\text{SKEL}} + \Delta\mathbf{J}_i$ and the posed joint locations $\mathbf{J}_{\mathbf{q}_i}^{\text{GT}}$ after the optimization and consider them as ground truth.

2.2 The SKEL-J model

The new SKEL-J model has the extended joint degree of freedom $\Delta\mathbf{J}$ and a new joint regressor $\mathcal{J}_{\text{SKEL-J}}$ learned from the computed registrations. Following SMPL [26], OSSO [15] and SKEL [17], the linear regressor takes as input the shaped vertices in T-pose $(\mathbf{T} + \boldsymbol{\beta}\mathbf{B})$ and predicts the new joint locations $\mathbf{J}^{\text{SKEL-J}}$ that best match the ground truth ones $\mathbf{J}_i^{\text{GT}}$. The regressor is learned with a non-negative least squares formulation and solved with an active set method [27].

3 Experiments

3.1 Bone registration accuracy

We start by evaluating how well the bone registrations $\mathbf{v}_{\text{skel}}^i$ capture the segmented point-cloud bones $\{\mathcal{B}_i^*\}$. We measure, for each subject, the mean point-to-mesh distances between the individual bones by computing Eq. 1 divided by the number of points in $\{\mathcal{B}_i^*\}$. In Fig. 2 left, we show examples of the computed $\mathbf{v}_{\text{skel}}^i$ registered to the bone pointclouds $\{\mathcal{B}_i^*\}$, and in Fig. 2 right, a boxplot of the mean distances for the full dataset, i.e. each sample is the mean distance over one subject. These faithfully capture the segmented individual bones with a mean accuracy below 6 mm.

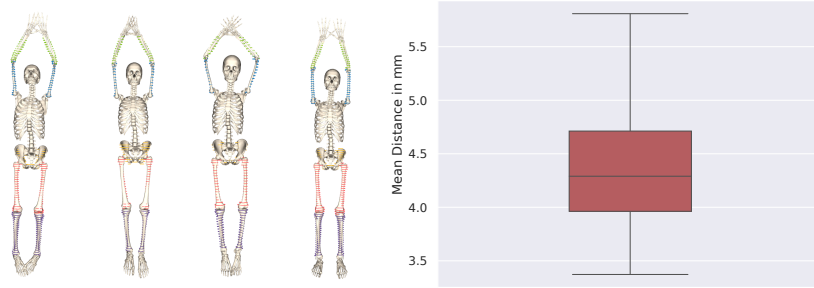


Fig. 2: Left: Bone pointclouds $\{\mathcal{B}_i^*\}$ (color coded) with registered $\mathbf{v}_{\text{skel}}^i$ (white). Right: Boxplot of the mean distances between $\{\mathcal{B}_i^*\}$ and $\mathbf{v}_{\text{skel}}^i$.

	Female				Male			
	Mean	Std	Min	Max	Mean	Std	Min	Max
OSSO	<u>9.79</u>	3.23	4.46	18.3	11.42	3.27	4.74	19.63
SKEL	9.95	3.34	5.02	22.74	10.67	3.06	5.09	19.59
SKEL-J	9.32	3.34	4.03	21.84	9.72	2.87	4.66	17.45

Table 1: Mean, standard deviation, min and max distances, in mm, between the test set bone ground truth point-clouds and the methods predicted meshes.

3.2 3D Bone prediction accuracy

Next, we compare the predicted 3D bones by OSSO [15], SKEL [17], HIT [21] and SKEL-J using the fitted SMPL meshes $\mathbf{S}$ as input. We could not compare to Wu et al. [13] as the code is not available. The evaluation is performed on test data (22 females, 14 males) that were not used to train OSSO, SKEL, HIT, or SKEL-J.

To evaluate the accuracy of the individual bone predictions, we compute the distance between the ground truth bone point-cloud and the predicted individual bone mesh. For each subject, we aggregate the distances for all individual bones. We assess statistical significance using the Wilcoxon signed-rank [28] test ($\alpha = 0.05$). Table 1 reports the statistics over all test subjects. Only SKEL-J and OSSO are not statistically different ($p > 0.05$) for females (underlined). SKEL-J consistently outperforms all other state of the art predictions for both genders ($p < 0.05$).

To get a finer per-individual bone analysis, Fig. 3 reports a boxplot of all the distances' distribution aggregated per individual bone over the test dataset, where each sample is one subject. Statistical significance ($p < 0.05$) is reported with a *. It is interesting to note that while SKEL-J outperforms SKEL, OSSO is more accurate for the Tibia and Fibula, both for males and females. On a general level, all these estimates still remain inaccurate when compared to the ground truth data, showing that there might still be room for improvement.

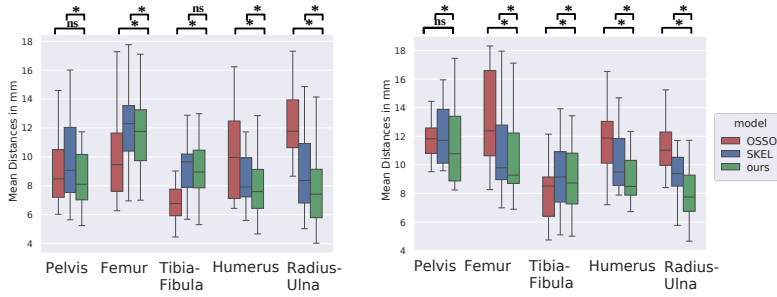


Fig. 3: Per individual bone quantitative comparison between OSSO, SKEL, and SKEL-J (ours). Left females, right males. Significance threshold fixed at 0.01.

Note that HIT only predicts a binary skeleton structure without the individual bone labels, thus, to compare HIT and SKEL-J we use a different metric: we compute the distance from each bone point-cloud vertex to the full predicted skeleton. For HIT we obtain mean distances of 11.54 mm for females and 8.83 mm for males, whereas for SKEL-J’s we obtain 8.86 mm for females and 9.35 mm for males. While for males both predictions are similar (HIT 0.5 mm better than SKEL-J with p-value of 0.03), other than the improved accuracy for females (p-value of $1e-8$), the main advantage of SKEL-J with respect to HIT is that individual bones are predicted. In Fig. 4 (and Sup. Mat. Fig. ??, Fig. ??) we present qualitative comparisons between all the predicted skeletons.

Evaluation on supine patients. In the used dataset HIT [21], all subjects are lying in a prone position. To start exploring how well the prediction methods perform in a different pose, we scanned 4 subjects in a supine position with the arms resting on the stomach. We used the same MRI sequence as in [29] but a different machine (3.0T - PRISMA Fit, Siemens Healthcare). The previous segmentation network was not able to properly segment this data due to the gap domain between the body positions and the machines. Thus we manually segmented the 4 cases. A visualization of the obtained SMPL fits and individual bone segmentations is shown in Sup. Mat. Fig. ?. We computed the same metrics as in the previous experiments and obtained an individual bone mean distance of 13.49 (std 8.74) for SKEL-J, 13.36 (std 9.31) for OSSO and 13.47 (std 8.68) for SKEL, as well as a binary bone mean distance of 14.42 (std 8.03) for HIT. In Fig. 5 we present qualitative results on two cases and in Sup. Mat. Fig. ?? the additional two subjects.

Most interestingly, all methods have a very similar performance in this case. As OSSO was learned from poses in a supine position with arms near the body, one could have expected it to exhibit an improved performance in this particular case. As visible in the qualitative results, this is not the case and the shoulder locations are consistently estimated too low with respect to the body for all methods. This raises awareness of the difficulty to generalize to unseen poses, in particular for the complex shoulder articulation.

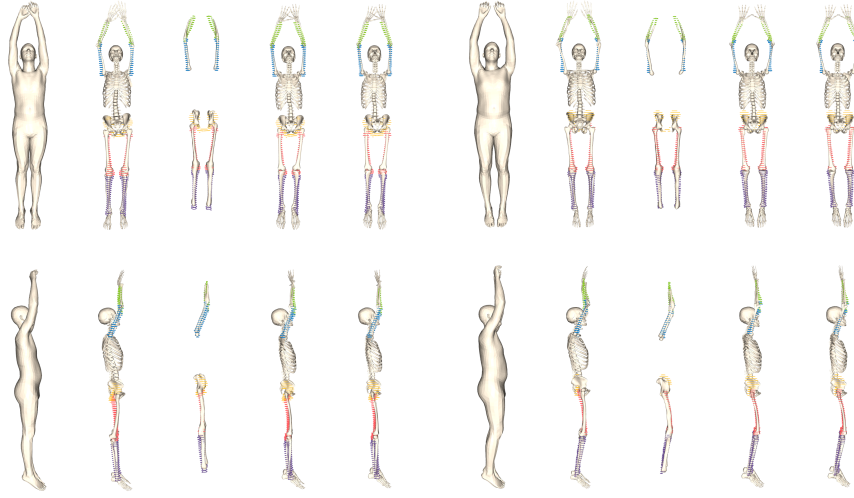


Fig. 4: For two subjects, from left to right: SMPL input, then the predictions by OSSO, HIT, SKEL and SKEL-J. Top: frontal view. Bottom: lateral view.

4 Conclusion

This paper tackles the problem of predicting the location of the bones inside the human body. One key aspect to solve this problem is to leverage clinical 3D data, where the bones and the skin are simultaneously visible. Towards that end, we leverage an existing dataset [21] that we enrich with individual bone segmentations. A second key aspect is the proper registration of this data to learn its correlations. To achieve accurate registrations, we extend the biomechanical model SKEL [17] into SKEL-J by including new degrees of freedom that allow to simultaneously capture the skin and the bones. From the registered data, we can learn to accurately predict the bone locations from the skin surface. The experiments show that SKEL-J outperforms the existing state of the art.

One limitation of our approach is the lack of variety of poses in the dataset. Our last experiment, in which the subjects lie in a supine position, never seen at training time, highlights the need to validate and improve the prediction accuracy for more and more diverse poses. Acquiring such 3D data with the existing medical devices is not straightforward, but different options, such as *standing CTs*, or Digital Motion X-Ray could provide rich information about the bone locations in different poses. Also, our work focuses on predicting the long bones and disregards important structures such as the spine, feet, hands, or the skull. Future work should further study how accurate the predictions of these structures can be.

Acknowledgments. We thank Kate Duquesne and Emmanuel Audenaert for the initial individual bone segmentations. Marilyn Keller was supported by

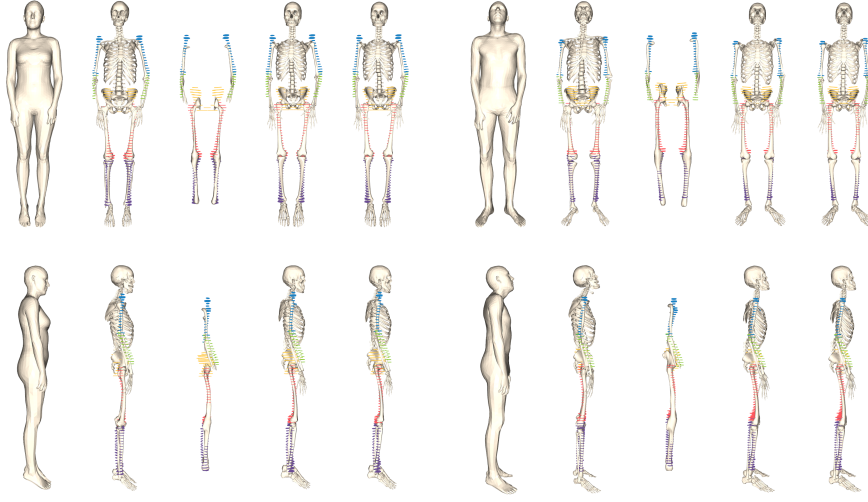


Fig. 5: From left to right: SMPL input in supine position, then the predictions by OSSO, HIT, SKEL and SKEL-J. Top: frontal view. Bottom: lateral view.

the International Max Planck Research School for Intelligent Systems. Léo Challer and Sergi Pujades' work was funded by the ANR JCJC SEMBA project.

Disclosure of Interests. MJB CoI Disclosure: https://files.is.tue.mpg.de/black/CoI_CVPR_2024.txt. The other authors have no competing interests to declare that are relevant to the content of this article.

References

1. Bauer, S., Wasza, J., Haase, S., Marosi, N., Hornegger, J.: Multi-modal surface registration for markerless initial patient setup in radiation therapy using microsoft's kinect sensor. In: 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops). pp. 1175–1181. IEEE (2011)
2. Karanam, S., Li, R., Yang, F., Hu, W., Chen, T., Wu, Z.: Towards contactless patient positioning. *IEEE transactions on medical imaging* **39**(8), 2701–2710 (2020)
3. Teixeira, B., Singh, V., Tamersoy, B., Prokein, A., Kapoor, A.: Automated ct lung cancer screening workflow using 3d camera. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 423–431. Springer (2023)
4. Gilles, B., Reveret, L., Pai, D.: Creating and animating subject-specific anatomical models **29**(8), 2340–2351 (Dec 2010). <https://doi.org/10.1111/j.1467-8659.2010.01718.x>
5. Ali-Hamadi, D., Liu, T., Gilles, B., Kavan, L., Faure, F., Palombi, O., Cani, M.P.: Anatomy transfer. *ACM Trans. Graphic.* **32**(6), 1–8 (Nov 2013). <https://doi.org/10.1145/2508363.2508415>, <https://doi.org/10.1145/2508363.2508415>
6. Saito, S., Zhou, Z.Y., Kavan, L.: Computational bodybuilding: Anatomically-based modeling of human bodies. *ACM Trans. Graphic.* **34**(4), 1–12 (2015)
7. Zhu, L., Hu, X., Kavan, L.: Adaptable anatomical models for realistic bone motion reconstruction **34**(2), 459–471 (2015). <https://doi.org/10.1111/cgf.12575>
8. Kadleček, P., Ichim, A.E., Liu, T., Krivánek, J., Kavan, L.: Reconstructing personalized anatomical models for physics-based body animation. *ACM Trans. Graphic.* **35**(6), 1–13 (Nov 2016). <https://doi.org/10.1145/2980179.2982438>, <https://doi.org/10.1145/2980179.2982438>
9. Ichim, A.E., Kadleček, P., Kavan, L., Pauly, M.: Phace: Physics-based face modeling and animation. *ACM Trans. Graphic.* **36**(4), 1–14 (2017). <https://doi.org/10.1145/3072959.3073664>
10. Shetty, K., Birkhold, A., Jaganathan, S., Strobel, N., Egger, B., Kowarschik, M., Maier, A.: Boss: Bones, organs and skin shape model. *Computers in Biology and Medicine* **165**, 107383 (2023). <https://doi.org/https://doi.org/10.1016/j.compbimed.2023.107383>, <https://www.sciencedirect.com/science/article/pii/S001048252300848X>
11. Li, J., Pepe, A., Luijten, G., Schwarz-Gsaxner, C., Kleesiek, J., Egger, J.: Anatomy completer: A multi-class completion framework for 3D anatomy reconstruction. In: International Workshop on Shape in Medical Imaging. pp. 1–14. Springer (2023)
12. Teixeira, B., Singh, V., Chen, T., Ma, K., Tamersoy, B., Wu, Y., Balashova, E., Comaniciu, D.: Generating synthetic x-ray images of a person from the surface geometry. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 9059–9067 (2018)
13. Wu, Y., Singh, V., Teixeira, B., Ma, K., Tamersoy, B., Krauss, A., Chen, T.: Towards generating personalized volumetric phantom from patient's surface geometry. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part I. pp. 171–179. Springer (2018)
14. Guo, H., Planche, B., Zheng, M., Karanam, S., Chen, T., Wu, Z.: SMPL-A: Modeling person-specific deformable anatomy. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 20814–20823 (2022)

15. Keller, M., Zuffi, S., Black, M.J., Pujades, S.: OSSO: Obtaining skeletal shape from outside. In: *Proceedings IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)*. pp. 20492–20501 (Jun 2022)
16. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott, P., Green, J., Landray, M., Liu, B., Matthews, P., Ong, G., Pell, J., Silman, A., Young, A., Sprosen, T., Peakman, T., Collins, R.: UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLOS Medicine* **12**(3), e1001779 (2015)
17. Keller, M., Werling, K., Shin, S., Delp, S., Pujades, S., C. Karen, L., Black, M.J.: From skin to skeleton: Towards biomechanically accurate 3D digital humans. In: *ACM ToG, Proc. SIGGRAPH Asia* (Dec 2023)
18. Mahmood, N., Ghorbani, N., Troje, N.F., Pons-Moll, G., Black, M.J.: AMASS: Archive of motion capture as surface shapes. pp. 5442–5451. *IEEE* (Oct 2019)
19. Delp, S.L., Anderson, F.C., Arnold, A.S., Loan, P., Habib, A., John, C.T., Guendelman, E., Thelen, D.G.: OpenSim: Open-source software to create and analyze dynamic simulations of movement. *IEEE Transactions on Biomedical Engineering* **54**(11), 1940–1950 (2007)
20. Werling, K., Raitor, M., Stingel, J., Hicks, J.L., Collins, S., Delp, S.L., Liu, C.K.: Rapid bilevel optimization to concurrently solve musculoskeletal scaling, marker registration, and inverse kinematic problems for human motion reconstruction. *bioRxiv* (2022). <https://doi.org/10.1101/2022.08.22.504896>, <https://www.biorxiv.org/content/early/2022/08/23/2022.08.22.504896>
21. Anonymous: Anonymous (2024)
22. Wenninger, S., Kemper, F., Schwanecke, U., Botsch, M.: Tailormer: Self-supervised learning of an anatomically constrained volumetric human shape model. In: *Computer Graphics Forum*. p. e15046. Wiley Online Library (2024)
23. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.: nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods* **18**, 203 – 211 (2020), <https://api.semanticscholar.org/CorpusID:227947847>
24. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., et al.: Segment anything. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. pp. 4015–4026 (2023)
25. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. *Nature Communications* **15**(1), 654 (2024)
26. Loper, M., Mahmood, N., Romero, J., Pons-Moll, G., Black, M.J.: Smpl: A skinned multi-person linear model. *ACM transactions on graphics (TOG)* **34**(6), 1–16 (2015)
27. Lawson, C.L., Hanson, R.J.: Solving least squares problems. *Society for Industrial and Applied Mathematics* (1995)
28. Wilcoxon, F.: Individual comparisons by ranking methods. *Biometrics Bulletin* **1**(6), 80–83 (1945), <http://www.jstor.org/stable/3001968>
29. Machann, J., Thamer, C., Schnoedt, B., Haap, M., Haring, H.U., Claussen, C.D., Stumvoll, M., Fritsche, A., Schick, F.: Standardized assessment of whole body adipose tissue topography by mri. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine* **21**(4), 455–462 (2005)