Mesh-based 3D Reconstruction from Bi-planar Radiographs

Moritz Jokeit[1,2] moritz.jokeit@hest.ethz.ch
[1] *Institute for Biomechanics, ETH Zurich, Zurich, Switzerland*
[2] *Department of Orthopaedics, Balgrist University Hospital, Zurich, Switzerland*

Ji Hyun Kim[1] jihyun.kim@hest.ethz.ch
Jess G. Snedeker[1,2] jess.snedeker@hest.ethz.ch
Mazda Farshad[2] mazda.farshad@balgrist.ch
Jonas Widmer[1,2] jonas.widmer@balgrist.ch

Editors: Under Review for MIDL 2022

Abstract

Reconstruction of 3D surfaces from sparse 2D data is a challenging problem that attracted increasing attention also in the medical field where image acquisition is expensive and the patients often bear high radiation doses (CT, fluoroscopy). Further, advances in computer-guided surgical assistant systems and preoperative planning necessitate fast 3D reconstruction from sparse image data. Recent learning-based approaches showed notable success in reconstructing primitive objects leveraging abundant artificial data sets. However, quality 3D data in the clinical context is often scarce. This motivates the exploitation of domain knowledge in form of anatomical shape priors to simplify the reconstruction problem. Further, mesh-sensitive applications (e.g., finite element analysis of implant design) greatly benefit from pre-defined mesh topologies. Thus, we propose a concept for implementing and training a learning-based patient-specific 3D reconstruction from bi-planar radiographs based on altering anatomical template meshes.

Keywords: 3D reconstruction, deep learning, mesh, medical imaging, X-ray, digitally reconstructed radiograph (DRR), style transfer, computed tomography (CT).

1. Introduction

Reconstruction and registration of the patient's anatomy become more important with advances in computer-assisted surgery and planning (Fürnstahl et al., 2012). Current approaches often rely on segmented CT scans, which are costly, time-consuming to acquire and expose the patient to high radiation doses. However, 3D reconstruction from scarce 2D data leads to an ill-posed problem that can not be solved with conventional methods. Here, learning-based solutions for sparse-view 3D reconstruction provide a promising alternative (Kasten et al., 2020). Wherever data are sparse, domain knowledge helps to simplify the solution process. Inspired by early medical segmentation algorithms using shape priors, we want to exploit anatomical model templates to enhance the learning-based 3D reconstruction from bi-planar radiographs. The end-to-end mesh prediction facilitates patient-specific musculoskeletal modeling, finite element analysis, or preoperative planning.

2. Methods

The proposed pipeline aims to reconstruct patient-specific geometries provided with only bi-planar radiographs and a template mesh of the desired anatomical structure. Figure 1 describes the workflow on the example of the pelvic bone. The core of our reconstruction pipeline is a deformation network that predicts a translation vector for each vertex of the mesh template which best fits the radiographic input. Following the idea of Wen et al. (2019), convolutional encoder networks (VGG-16) extract information from the two radiographs with bi-planar perceptual feature pooling. The resulting feature vector is then concatenated with the first layer of the reconstruction network. As a primary variant, we suggest employing a fully-connected architecture as in Pan et al. (2019) with ReLU activations in the first three layers and tanh activations in the output layer. Finally, applying the predicted translation vectors to the template nodes yields the patient-specific mesh.

To train the network we propose a loss term that simply considers the sum of the euclidean distances between every vertex x_i and y_i of the predicted mesh and the ground truth, respectively: $\mathcal{L}_d = \sum_i ||x_i - y_i||^2$. Extending the loss function with geometry regularizers potentially benefits the mesh's smoothness (Wen et al., 2019).

To obtain training data, we derive digitally reconstructed radiographs (DRRs) from 1,184 CT volumes (Liu et al., 2021). Then, an iterative closest point algorithm fits the template mesh to the corresponding 3D segmentations generating the ground truth mesh that complements the DRR (e.g., Scalismo). A frequent problem of DRRs is their mismatch in style regarding the emulated image modality. However, our preliminary results showed that generative networks like CycleGAN can be successfully trained to imitate the desired image style in an unsupervised manner.

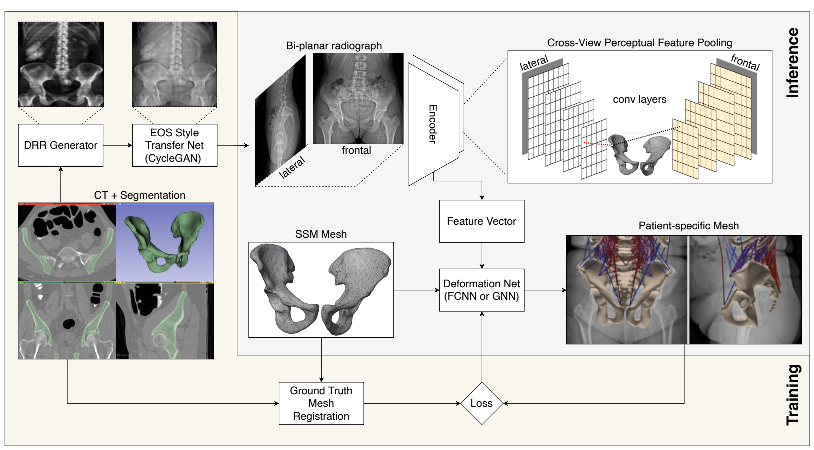

Figure 1: Concept for the implementation of the proposed reconstruction pipeline. Cross-view perceptual feature pooling adapted from Wen et al. (2019).

3. Conclusion

We believe combining shape priors and learning-based reconstruction has potential in a plethora of medical applications, e.g., a predefined mesh facilitates further use in patient-specific models and implant design. However, a mesh-based reconstruction comes not without challenges regarding the accuracy and implementational complexity. The presented concept will likely adapt to new challenges and ideas that occur during the implementation process. For instance, Wen et al. (2019) already demonstrated that graph convolutions are well suited for mesh-based reconstruction. The known geometry and physical properties of the X-ray device offer even more potential for domain knowledge exploitation. Further, the image encoder could be replaced by lighter and more efficient network architectures.

Acknowledgments

We thank Prof. Dr. Philipp Fürnstahl and Dr. Hooman Esfandiari from ROCS, Balgrist University Hospital, University of Zurich, for their valuable inputs, and the Promedica Stiftung, Chur, for financial support. This research was supported by grants from NVIDIA and utilized NVIDIA RTX A6000 GPUs.

References

Philipp Fürnstahl, Gábor Székely, Christian Gerber, Jürg Hodler, Jess Gerrit Snedeker, and Matthias Harders. Computer assisted reconstruction of complex proximal humerus fractures for preoperative planning. *Medical Image Analysis*, 16(3):704–720, April 2012. ISSN 1361-8415. doi: 10.1016/j.media.2010.07.012. URL https://www.sciencedirect.com/science/article/pii/S1361841510001040.

Yoni Kasten, Daniel Doktofsky, and Ilya Kovler. End-To-End Convolutional Neural Network for 3D Reconstruction of Knee Bones from Bi-planar X-Ray Images. In Farah Deeba, Patricia Johnson, Tobias Würfl, and Jong Chul Ye, editors, *Machine Learning for Medical Image Reconstruction*, Lecture Notes in Computer Science, pages 123–133, Cham, 2020. Springer International Publishing. ISBN 9783030615987. doi: 10.1007/978-3-030-61598-7_12.

Pengbo Liu, Hu Han, Yuanqi Du, Heqin Zhu, Yinhao Li, Feng Gu, Honghu Xiao, Jun Li, Chunpeng Zhao, Li Xiao, Xinbao Wu, and S. Kevin Zhou. Deep Learning to Segment Pelvic Bones: Large-scale CT Datasets and Baseline Models. *arXiv:2012.08721 [cs]*, March 2021. URL http://arxiv.org/abs/2012.08721. arXiv: 2012.08721.

Junyi Pan, Xiaoguang Han, Weikai Chen, Jiapeng Tang, and Kui Jia. Deep Mesh Reconstruction From Single RGB Images via Topology Modification Networks. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 9963–9972, Seoul, Korea (South), October 2019. IEEE. ISBN 978-1-72814-803-8. doi: 10.1109/ICCV.2019.01006. URL https://ieeexplore.ieee.org/document/9009447/.

Chao Wen, Yinda Zhang, Zhuwen Li, and Yanwei Fu. Pixel2Mesh++: Multi-View 3D Mesh Generation via Deformation. *arXiv:1908.01491 [cs]*, August 2019. URL http://arxiv.org/abs/1908.01491. arXiv: 1908.01491.