Maximilian Eibl, Martin Gaedke. (Hrsg.): INFORMATIK 2017,
Lecture Notes in Informatics (LNI), Gesellschaft für Informatik, Bonn 2017 891

A Dense Statistical Model of Facial Soft Tissue Thickness

Thomas Gietzen[1], Robert Brylka[1], Ulrich Schwanecke[1] and Elmar Schömer[2]

Abstract: Ambient intelligence become more and more ubiquitous and help people achieving a more natural interaction with their electronically enhanced environment. One vital natural interface between humans and ambient intelligence are embodied conversational agents. Thereby, the acceptance of these virtual characters is all the greater, the more natural they look and behave. Since humans pay particular attention to the face, a natural-looking animation of the face is very important. In this paper we present a dense statistical model of facial soft tissue thickness that can be used to build accurate physics-based facial animations. The presented model not only can help to generate more natural facial animations of virtual characters but also can be used in other research domains such as forensic anthropology or medicine. Especially in the field of dentistry and orthodontics in particularly younger people and children are increasingly examined using X-ray technology. Thereby more and more volumetric images are generated, which further increase cost as well as the induced radiation dose. Here, for example, our statistical model can provide the basis for a new volumetric reconstruction process of a human's facial bones in a cost-effective manner and with low radiation exposure.

Keywords: 3D volume registration; ambient intelligence; conversational agents; facial soft tissue thickness; lateral cephalogram; medical imaging; statistical model

1 Introduction

Facial soft tissue thickness (FSTT) is determined by measuring the distance between the skin surface and the underlying skull bones. FSTT plays an important role in medical diagnosis and therapy. A thorough review of soft tissue depth studies and their applications is given by [SS08]. Soft tissue depth studies are based on a variety of measurement techniques such as needle puncture, ultrasound, computed tomography (CT) or magnetic resonance imaging (MRI). Thereby, besides the different measurement technics the variation of the direction in which FSTT is measured (based on the normals on the bones or skin surfaces or oriented towards other soft tissue features) have led to greatly varying results which are hard to compare. Additionally, most medical studies determine and use FSTT only at various predefined medical landmarks.

Beside medicine, statistics about FSTT can be used in forensic anthropology for craniofacial reconstruction, i.e. for reconstructing the morphology of a particular face, if only a skull is

[1] RheinMain University of Applied Sciences, Design Computer Science Media, Unter den Eichen 5, 65195 Wiesbaden, firstname.surname@hs-rm.de
[2] Johannes Gutenberg-University Mainz, Institute of Computer Science, Staudingerweg 9, 55099 Mainz, schoemer@informatik.uni-mainz.de

doi:10.18420/in2017_90

given or in physics-based facial animation of virtual characters that can serve as embodied conversational agents. Thereby, it is important to have a dense map of the FSTT and not only information about some predefined medical landmarks. Further, a dense statistical model of the FSTT is used to determining a coarse volumetric reconstruction of the human's facial bones having only an optical face scan as input.

In this paper we present an approach for the automatic generation of a dense statistical model of the FSTT based on CT data of the human head. We generate the statistical model by registering a volumetric template model of the skull[3] into each CT data set and determine the FSTT for any vertex of the template skull. Previous approaches to register a template skull into CT data, such as [SK15] concentrate on aligning the upper and back part of the skull (neurocranium). They are not appropriate to register the fine structure of the facial bones. Our main contribution is a dense statistical model of FSTT based on a new method for fine registration of a template skull to the facial bones of a skull extracted from CT data.

2 Generating a volumetric model from a single lateral cephalogram

Beside the generation of physics-based facial animations of virtual characters, we are interested in a new method for determine the exact surface of the facial bones with minimal radiation exposure. Thereby, we will use three main inputs (see fig. 1): a) a single lateral cephalogram, b) an optical face scan, and c) the statistical model of the FSTT.

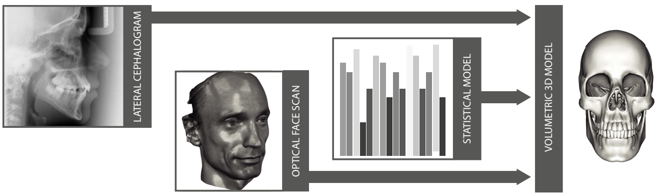


Fig. 1: Low radiation exposure based volumetric reconstruction process.

A lateral cephalogram (LC) is an X-ray taken from the side of a patient. It is usually used to determine the relationship between top and bottom jaw (maxilla and mandible), i.e. to assess the nature of a patient's bite. A LC has only minimal perspective distortions and therefore contains suitable information about the soft tissue thickness on the facial midline.

Optical face scanning provides an exact 3D model of the skin surface. While various commercial solutions such as *Agisoft PhotoScan* or *Autodesk Recap* exist to generate 3D spatial data from a variety of digital images, we use a self-developed passive scanning system tailored to the requirements of our reconstruction process [Gr17].

The two measurements described above (lateral cephalogram and face scan) are not sufficient to reconstruct the facial skull. Thus, in order to reconstruct the facial skull, we will later combine both information using the dense statistical model of FSTT presented in this paper.

[3] Based on the model `www.turbosquid.com/3d-models/3d-human-skull/691781`

3 Generating a dense statistical model

The construction of the statistical model is based on CT data sets of the human head. An overview of the data processing pipeline is depicted in figure 2. It starts with a preprocessing step normalizing the information in the CT data and extracting triangular meshes of the skin surface and the skull. Next, a volumetric template skull is registered into this data. This results in a fine registered template skull. Finally, the mean and standard deviation of the FSTT of all vertices of the template skull based on the Hausdorff distance between the deformed template skulls and each of the extracted skin surfaces is determined.

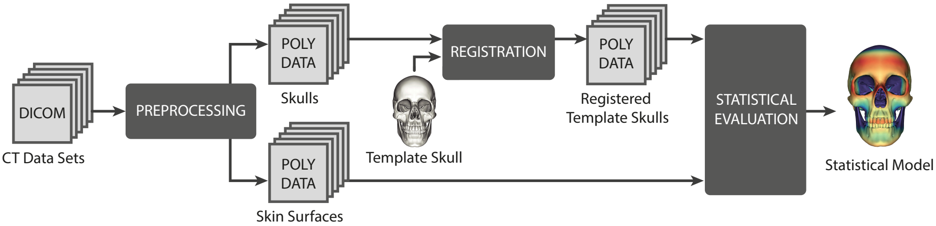


Fig. 2: Processing pipeline for generating the statistical model of FSTT.

3.1 Preprocessing the CT data sets

To extract exact polygonal meshes of the skin surface and the skull from a given CT data set of the head a preprocessing step is needed. In this step the data is cropped below the mandibular bone. In order to determine the CT slice containing the most inferior point of the mandibular (see figure 3) we first smooth each slice and then generate a binary image with a threshold representing bones in Hounsfield units. Next, we choose the slice s_j with $j = \arg\max_{i=2,\ldots,\#slices} |y_i - y_{i-1}|$ where y_i is the number of the first non-zero row in slice number i and add an offset to include the skin surface of the chin.

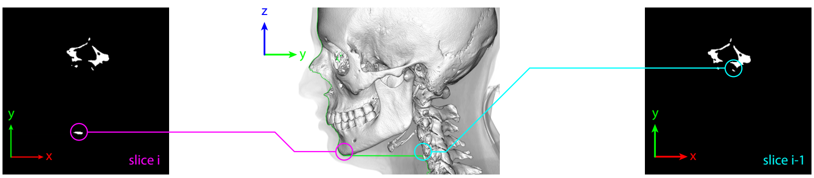


Fig. 3: Determination of the cropping position based on detection of the mandibular bone.

Once we found the cutting slice and cropped the volumetric data, we extract polygonal meshes of the skin surface and the skull using the Marching Cubes algorithm [LC87]. The skull mesh is cleaned from unwanted parts such as the spin using a connectivity filter with seed points being the first appearance of bone structure in the CT data sets representing the cranium and the found point on the mandibular bone. Finally all skull meshes are decimated to a common point density, surface normals are generated and internal bone structures are removed by determining all vertices that are not visible from the outside of the skull.

3.2 Template skull registration

The registration of the template skull to an extracted skull consists of two main stages: 1. an initial rigid transformation of an extracted skull to the template skull and 2. an elastic transformation in which we deform the template skull into the extracted one.

Initial alignment

The initial alignment of the template skull with the CT data starts with a coarse registration step using the Fast Global Registration (FGR) algorithm presented in [ZPK16]. We choose this algorithm as a global alignment approach because it does not need an initialization step and works well on noisy and only partially overlapping 3D surfaces.

Figure 4a shows the template skull (white) and an extracted skull (cyan) in initial position (defined by the CT) and the correspondences between the candidate matches. As depicted in figure 4b the FGR algorithm provides already a very good alignment. Nevertheless, we perform a second alignment step applying a rigid transformation with standard ICP, which results in a significant improvement of the alignment quality as shown in figure 4c.

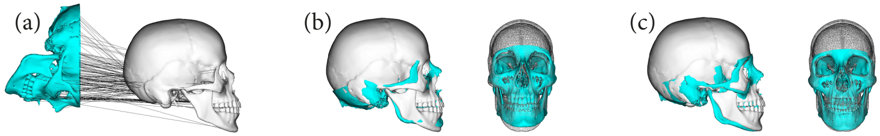


Fig. 4: Initial alignment steps: a) original configuration given by CT position, b) registration with FGR algorithm, c) ICP based refinement.

Deformation

In the second stage we perform a fine registration of the template skull into the extracted skull. Therefore, we morph our template skull into the extracted skull using ShapeOp [De15]. ShapeOp is a physics solver that integrates a variety of constraints especially projective constraints, dynamics and handle-based shape space exploration. ShapeOp can be used with a variety of geometry representations as point clouds, polygon meshes or tetrahedral meshes. To simulate the volume of the skull bones we use a tetrahedral mesh representation of our template skull as an input to the solver. The deformation steps based on weighted correspondences between points on the tetrahedral mesh and points which are on or close to the extracted skull. The weighting determines how strongly a point correspondence effects the deformation. Figure 5 shows our deformation pipeline, which consists of several deformation steps with different approaches to find point correspondences.

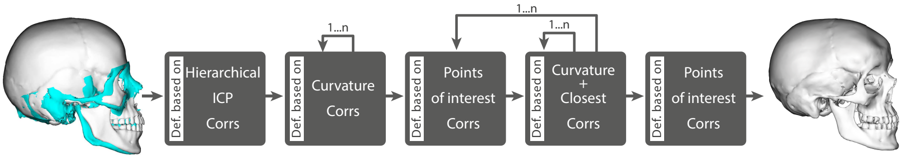
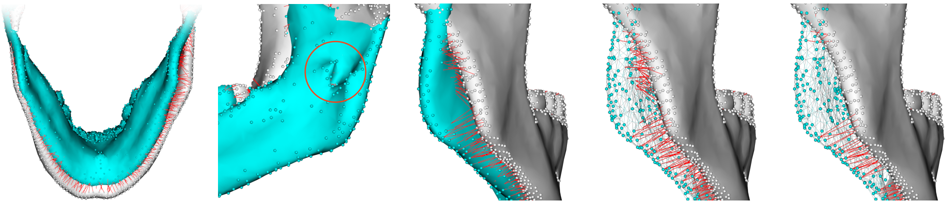


Fig. 5: Deformation pipeline with various approaches for correspondence search.

Closest (curvature) point correspondences. Point correspondences between meshes which are already close to each other can be determined based on the distance between points. Especially for complex shapes such as the facial skull this approach often leads to false correspondences. We therefore extend the correspondence search with additional constraints depending on the deformation step. Generally speaking, we restrict the search by a maximum distance and a minimum normal deviation constraint.

In the primary deformation steps where the meshes are far apart we additionally restrict the candidates for closest correspondence to points with high curvature value (closest curvature correspondences). Figure 6a shows the correspondence matches for curvature points (for the sake of clarity only the mandible is displayed). Nevertheless these constraints do not prevent wrong correspondence matches as shown in 6c and 6d. In addition, we check any estimated closest curvature correspondence match for their similarity given by a Fast Point Feature Histogram [RBB09] (FPFH). A FPFH encodes geometrical properties in a point k-neighborhood by generalizing the mean curvature around this point and provide more information than the normals deviation check. With this constraint it is possible to limit wrong correspondences as visualized in 6e.



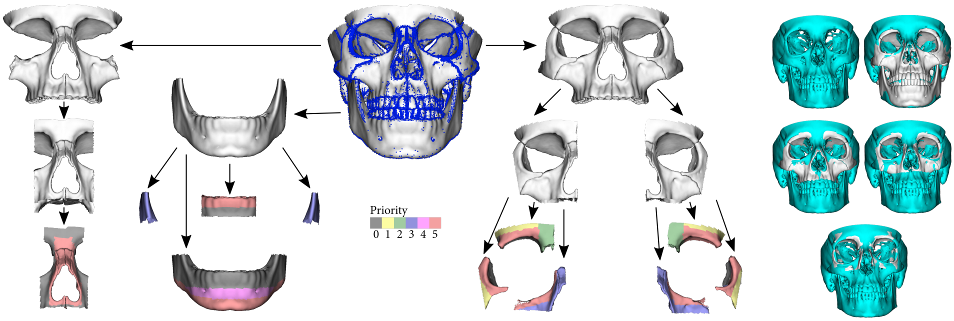
(a) (b) (c) (d) (e)

Fig. 6: Closest curvature correspondences: a) example of correspondence matches for points with high curvature value; b),c),d) correspondence mismatch on a challenging part as the mandible foramen; e) correspondence mismatch elimination with a FPFH similarity check.

Unfortunately, after initial alignment some parts of the template skull are still too far away from an extracted skull for a closest point correspondence search which hampers a good deformation. Thus, we introduce a new hierarchical approach to reliably determine point correspondences.

Hierarchical ICP correspondences. In order to find more correct correspondences for the first deformation step we introduce a hierarchical ICP approach fitting individual parts of the template skull. As shown in 7a (middle) we use the template skull without the skullcap (calvaria) since most of the CT data sets do not include this part. We register the hierarchically divided parts into the extracted skull using a similarity transformation. Since the smallest parts in our hierarchy might include overlapping sections we additionally provide a priority list for any point and use only the points with highest priority. Figure 7a shows a few examples of the hierarchical divided parts and demonstrates the priority coding for the five priorities going from yellow for low priority up to red for high priority. Dark gray marks the additional priority zero which excludes points from being considered in the deformation step at all.

The priority of any point is used as a weight in the following deformation step. Figure 7b shows an example of the result for the hierarchical ICP fitting the orbit parts. Since the topology between the tetrahedral representation and the template skull parts is not the same we compute the correspondences using the closest point constraint, considering only points whose curvature value is greater than a defined threshold (figure 7a blue colored points). Based on these correspondences we perform the first deformation step. After this step we continue with an additional closest curvature correspondence search between template skull and extracted skull and subsequent deformations based on those correspondences.



(a) (b)

Fig. 7: Hierarchical ICP: a) overview of the facial skull bones hierarchy used by the correspondence search, b) example of the hierarchical registration for the orbit parts.

Points of interest correspondences. Even the described FPFH constraint can not prevent for inaccurate correspondence matches which lead to excessive mesh distortion after a few deformation steps (see figure 8a, b). To prevent this shape degradation we introduce a deformation step based on predefined points of interest (see figure 8). After one deformation step with our hierarchical ICP approach and several deformation steps based on closest curvature correspondences we use our undeformed volumetric template skull and deform it into the current interim result. Figure 8d shows a close view of the mandible part and the correspondence matches based on the points of interest. Figure 8e depicts the result of the deformation step based on this approach. With this deformation step we get a deformed template which is a little less aligned to the extracted skull but much smoother than it was after the deformation step before (see figure 8f). We perform this deformation step several times during the deformation pipeline considering more and more points of interest in every iteration as colored encode in figure 8c and finally use all points in the last deformation step.

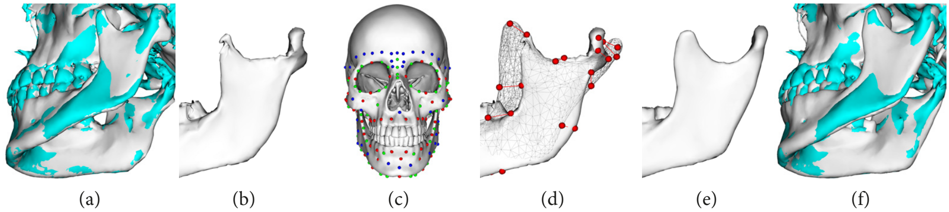


Fig. 8: Points of interest correspondences: a) interim result of template registration, b) mesh distortion at the mandibular notch, c) points of interest used for correspondence search, d) correspondence matches between an undeformed template (wireframe) and a distorted one, e), f) result of the deformation based on points of interest.

Dense statistical model

In order to determinate the FSTT we calculate the Hausdorff distance between the deformed template skull and the extracted skin surface for each data set. We create the statistical model by computing a mean skull based on all deformed skulls and determine the mean and the standard deviation of the Hausdorff distances for any point of the deformed template skulls. Currently, our statistical model is based on 40 different Caucasians (23 males and 17 females) with a mean age of 28. The age of the subjects vary between 17 and 35 years[4].

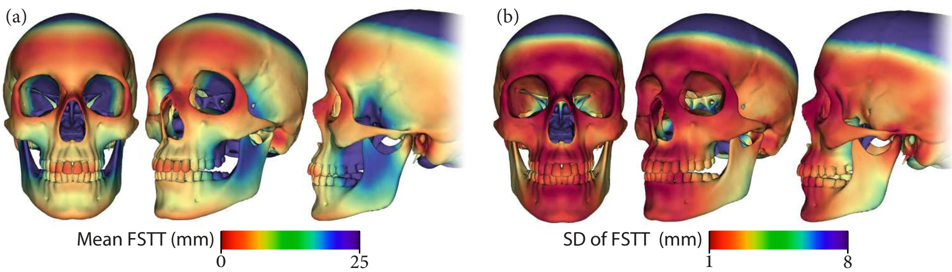


Fig. 9: Statistical model of the FSTT on a mean skull: a) Mean FSTT, b) Standard deviation.

A skull with color encoded mean and standard deviation of soft tissue thicknesses is depicted in figure 9. The statistics for predefined landmarks used in literature is shown in figure 10.

[4] This work is funded by the Federal Ministry of Education and Research (BMBF), grant No. IngenieurNachwuchs 03FH029IX5. We gratefully acknowledge the Department of Diagnostic and Interventional Radiology, University Medical Center of the Johannes Gutenberg University Mainz, Germany for providing us with the DICOM-data.

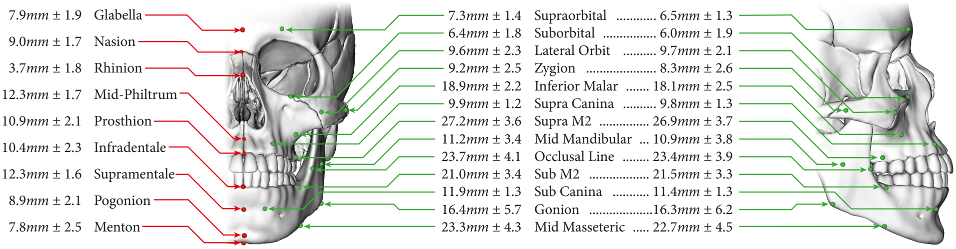


Fig. 10: Mean and standard deviation values of FSTT on predefined medical landmarks. Note that all measurements fall into the range presented in [SS08].

4 Conclusion and Future Work

We presented a method to generate a dense statistical model of the FSTT based on CT data and a volumetric skull template. Thereby, the values obtained on predefined medical landmarks (see figure 10) match the measurements presented in [SS08]. Next, together with medical experts, we will carefully examine the accuracy of our model. We will also use the statistics to gain more accurate physics-based facial animations and to combine a single lateral cephalogram and a face scan to obtain a volumetric reconstruction of the facial skull as described in section 2.

References

[De15] Deuss, Mario; Deleuran, Anders Holden; Bouaziz, Sofien; Deng, Bailin; Piker, Daniel; ; Pauly, Mark: ShapeOp—A Robust and Extensible Geometric Modelling Paradigm. Springer International Publishing, Cham, 2015.

[Gr17] Gruenewald, Torben: Multikamera Gesichtsrekonstruktion. Master's thesis, RheinMain University of Applied Sciences, 2017.

[LC87] Lorensen, William E.; Cline, Harvey E.: Marching Cubes: A High Resolution 3D Surface Construction Algorithm. SIGGRAPH '87, ACM, New York, NY, USA, pp. 163–169, 1987.

[RBB09] Rusu, Radu Bogdan; Blodow, Nico; Beetz, Michael: Fast Point Feature Histograms (FPFH) for 3D Registration. In: Proceedings of the 2009 IEEE International Conference on Robotics and Automation. ICRA'09, IEEE Press, Piscataway, NJ, USA, pp. 1848–1853, 2009.

[SK15] Sahillioğlu, Yusuf; Kavan, Ladislav: Skuller: A volumetric shape registration algorithm for modelling skull deformities. Medical image analysis, 23(1):15–27, 2015.

[SS08] Stephan, Carl N.; Simpson, Ellie K.: Facial Soft Tissue Depths in Craniofacial Identification (Part I): An Analytical Review of the Published Adult Data. Journal of Forensic Sciences, 78(6):1257–1272, 2008.

[ZPK16] Zhou, Qian-Yi; Park, Jaesik; Koltun, Vladlen: Fast Global Registration. In: ECCV 2016. pp. 766–782, 2016.