Building Accurate Physics-based Face Models from Data Supplemental Material

PETR KADLEČEK, University of Utah, USA and Charles University, Czech Republic
LADISLAV KAVAN, University of Utah, USA

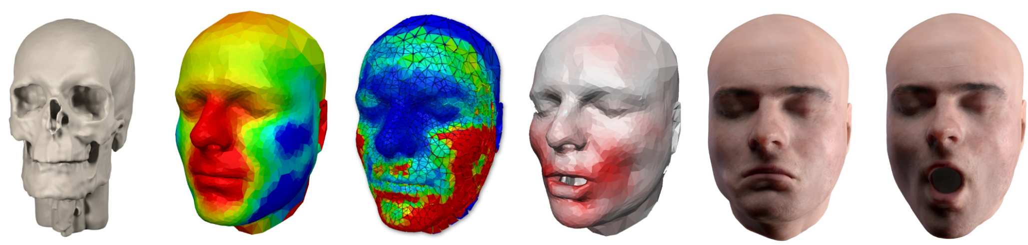

Fig. 1. We build a static anatomical face model from the MRI and use 3D surface scans as training data to learn mechanical parameters that explain deformations of the real face using physics-based simulation.

The human face is an anatomical system exhibiting heterogenous and anisotropic mechanical behavior. This leads to complex deformations even in a neutral facial expression due to external forces such as gravity. We start by building a volumetric model from magnetic resonance images of a neutral facial expression. To obtain data on facial deformations we capture and register 3D scans of the face with different gravity directions and with various facial expressions. Our main contribution consists in solving an inverse physics problem where we learn mechanical properties of the face from our training data (3D scans). Specifically, we learn heterogenous stiffness and prestrain (which introduces anisotropy). The generalization capability of our resulting physics-based model is tested on 3D scans. We demonstrate that our model generates predictions of facial deformations more accurately than recent related physics-based techniques.

Additional Key Words and Phrases: physics-based simulation, anatomical modeling, facial animation

1 DATA CAPTURE AND PROCESSING

We started by capturing four MRI scans of a neutral expression of one subject derived from two sequences (MP2RAGE INV1/INV2/UNI and T2, see Figure 3) and processed them using 3D Slicer [Fedorov et al. 2012]. Specifically, we applied Bias Field Correction to eliminate intensity variations and then denoised the data using Gradient Anisotropic Diffusion filters. We proceeded with segmentation of soft and hard tissues using thresholding and region growing. Note that segmentation of the skull and the mandible from MRI is a challenging task because bone tissue does not produce much signal and it is therefore easy to confuse with air, which can be problematic in areas such as the sinuses.

To address these challenges we tried modern MRI sequences better suited for solids, specifically, ultrashort echo time (UTE) [Robson and Bydder 2006]. The signal strength on bone tissue improved, but the images were very blurry. Therefore, we instead applied a "pseudo-CT" approach [Torrado-Carvajal et al. 2016] which is a data-driven technique to convert MR images into CT-like images. Our resulting segmentations of the bones and skin can be seen in Figure 2.

Authors' addresses: Petr Kadleček, University of Utah, Salt Lake City, UT, USA, Charles University, Prague, Czech Republic, petr.kadlecek@gmail.com; Ladislav Kavan, University of Utah, Salt Lake City, UT, USA, ladislav.kavan@gmail.com.

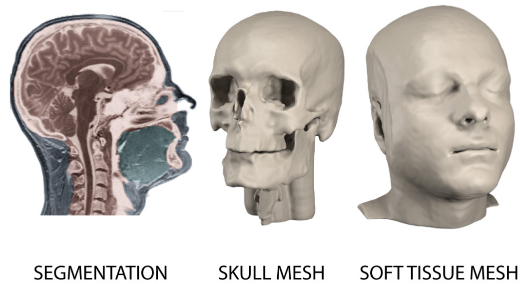

Fig. 2. Segmentation of soft tissues and the bones from our MRI scans.

When finished with 3D image processing, we used the marching cubes algorithm [Lorensen and Cline 1987] to create triangle meshes corresponding to the outer surface of the head (the skin) and the outer surfaces of the bones (the skull and the mandible). The triangle meshes were re-meshed using Meshmixer [Schmidt and Singh 2010] to improve mesh quality and used as constraints in TetGen [Si 2015] which produced our final tet-meshes, see Figure 4.

1.1 Surface 3D scans

Even though modern MRI scanners are very powerful and provide good resolution volumetric scans of the human body without any radiation, they have limitations. In addition to the high cost, long scanning times and limited availability of MRI machines, an important limitation is that the subject must remain motionless inside the MRI scanner for several minutes (depending on the sequence). This means that MRI scanning of most facial expressions is practically impossible because muscle fatigue would prevent the subject from staying still for minutes. Important advances in Real-time MRI have been made in recent years [Zhang et al. 2014], however, the methods are generally limited

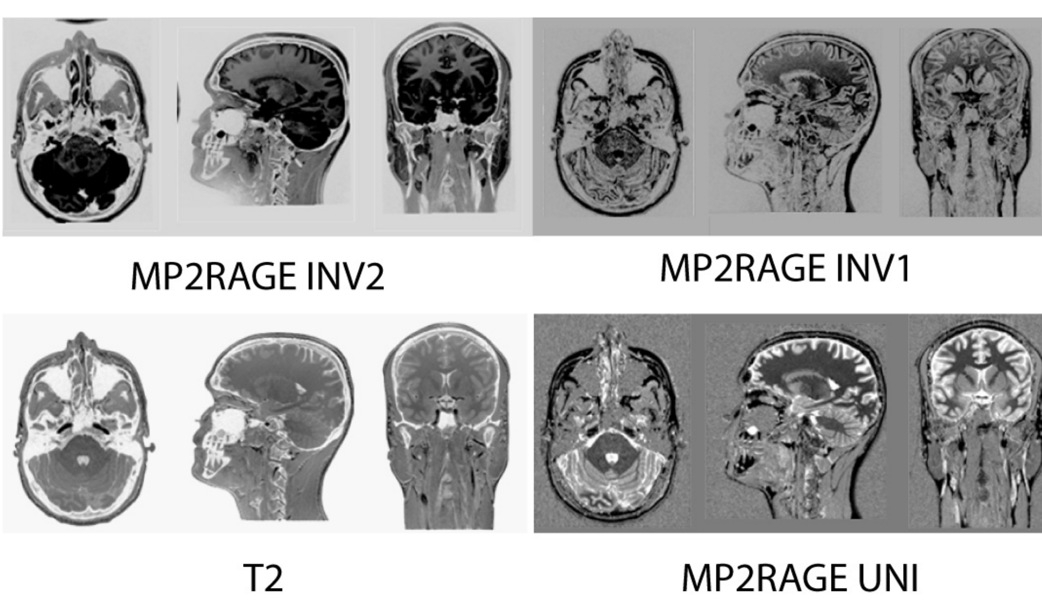

Fig. 3. Examples of slices from our MRI sequences used to build our model.

to either 1) very low-resolution volumetric imaging or 2) high-resolution single-slice images; neither is adequate for our purposes.

Instead, we captured geometry of deformed facial shapes using a structured light scanner (Artec Spider), producing detailed 3D scans of the skin. Specifically, in our deformed facial shapes we vary gravity directions (by changing the subject's head orientation) and facial expressions (by asking the subject to smile, frown, etc.). The face is quite supple and varying gravity directions results in surprisingly large skin displacements.

Registration. A necessary pre-requisite for subsequent automatic parameter fitting is registration: finding correspondences between the captured 3D scans. To do this, we first register the tet-mesh we built from the MRI scan with a 3D scan of a neutral expression in the supine pose, i.e., similar setup as in the MRI scanner where the subject is reclining on a motorized patient table. The two shapes (MRI and 3D scan) are close, but not identical due to geometric distortions of MRI images. Geometric distortion is a well-studied problem [Baldwin et al. 2007] and its corrections are standard, however, despite all efforts some geometric distortion still remains. We deform our tet-mesh to match the 3D scan using non-rigid Iterative Closest Points (ICP), enabling small deformations of the entire mesh including the bones to compensate for the geometric distortions of the MRI. When finished, the surface of the tet-mesh is closely aligned with the 3D scan, allowing us to transfer the albedo map from the 3D scan to the surface of our tet-mesh. The albedo map with painted markers (see Figure 5) helps us to find correspondences among our set of 3D scans of deformed facial shapes. Specifically, we apply another non-rigid ICP process with normalized cross-correlation of image patches as a data term [Beeler et al. 2011] and volumetric deformation of the tet-mesh as a regularization term, this time assuming the skull and the jaw are rigid, because the shapes of the bones must be the same in all physiologically deformed facial shapes. Finding the rigid transformation of the skull in each of the shapes gives us a "rigid pre-stabilization", i.e., an estimate of compensation for global head motion [Beeler and Bradley 2014]. This is only an initial guess (hence the term "pre-stabilization") which is refined when solving our inverse modeling problem. The output of this phase is a tet-mesh discretizing facial soft tissues which is accurately registered with our 3D scans of facial deformations. These data will be used to train our mechanical

High resolution mesh

# vertices:	5 441
# tetrahedra:	20 896
# faces:	8 754

Medium resolution mesh

# vertices:	2 698
# tetrahedra:	10 538
# faces:	4 218

Low resolution mesh

# vertices:	1 268
# tetrahedra:	4 445
# faces:	2 334

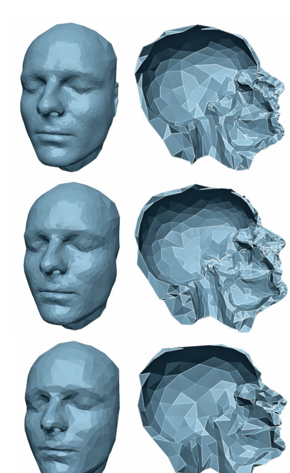

Fig. 4. Our tet-meshes with varying resolution.

Fig. 5. Registration of textured 3D scans.

model of the face. We prepared three versions of our tet-meshes with varying resolutions (Figure 4) which proved to be very helpful for experimentation and debugging.

REFERENCES

Lesley N Baldwin, Keith Wachowicz, Steven D Thomas, Ryan Rivest, and B Gino Fallone. 2007. Characterization, prediction, and correction of geometric distortion in 3T MR images. *Medical Physics* 34, 2 (2007), 388–399.

Thabo Beeler and Derek Bradley. 2014. Rigid stabilization of facial expressions. *ACM Trans. Graph.* 33, 4 (2014), 44.

Thabo Beeler, Fabian Hahn, Derek Bradley, Bernd Bickel, Paul Beardsley, Craig Gotsman, Robert W Sumner, and Markus Gross. 2011. High-quality passive facial performance capture using anchor frames. In *ACM Transactions on Graphics (TOG)*, Vol. 30. ACM, 75.

Andriy Fedorov, Reinhard Beichel, Jayashree Kalpathy-Cramer, Julien Finet, Jean-Christophe Fillion-Robin, Sonia Pujol, Christian Bauer, Dominique Jennings, Fiona Fennessy, Milan Sonka, et al. 2012. 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magnetic resonance imaging* 30, 9 (2012), 1323–1341.

William E Lorensen and Harvey E Cline. 1987. Marching cubes: A high resolution 3D surface construction algorithm. In *ACM siggraph computer graphics*, Vol. 21. ACM, 163–169.

Matthew D Robson and Graeme M Bydder. 2006. Clinical ultrashort echo time imaging of bone and other connective tissues. *NMR in Biomedicine* 19, 7 (2006), 765–780.

Ryan Schmidt and Karan Singh. 2010. Meshmixer: an interface for rapid mesh composition. In *ACM SIGGRAPH 2010 Talks*. ACM, 6.

Hang Si. 2015. TetGen, a Delaunay-based quality tetrahedral mesh generator. *ACM Transactions on Mathematical Software (TOMS)* 41, 2 (2015), 11.

Angel Torrado-Carvajal, Joaquin L Herraiz, Eduardo Alcain, Antonio S Montemayor, Lina Garcia-Cañamaque, Juan A Hernandez-Tamames, Yves Rozenholc, and Norberto Malpica. 2016. Fast patch-based pseudo-CT synthesis from T1-weighted MR images for PET/MR attenuation correction in brain studies. *Journal of Nuclear Medicine* 57, 1 (2016), 136–143.

Shuo Zhang, Arun A Joseph, Dirk Voit, Sebastian Schaetz, Klaus-Dietmar Merboldt, Christina Unterberg-Buchwald, Anja Hennemuth, Joachim Lotz, and Jens Frahm. 2014. Real-time magnetic resonance imaging of cardiac function and flowâĂŤrecent progress. *Quantitative imaging in medicine and surgery* 4, 5 (2014), 313.