Algorithmic Surface Extraction from MRI Data
Modelling the Human Vocal Tract

D. Aalto[1,2], J. Helle[3], A. Huhtala[4], A. Kivelä[4], J. Malinen[4], J. Saunavaara[5] and T. Ronkka[6]

[1]*Inst. of Behavioural Sciences (SigMe Group), University of Helsinki, Helsinki, Finland*

[2]*Dept. of Signal Processing and Acoustics, Aalto University, P.O. BOX 13000, FI-00076 Aalto, Finland*

[3]*Media Factory, Aalto University, P.O. BOX 31000, FI-00076 Aalto, Finland*

[4]*Dept. of Mathematics and Systems Analysis, Aalto University, P.O. BOX 11100, FI-00076 Aalto, Finland*

[5]*Dept. of Radiology, Medical Imaging Centre of Southwest Finland, University of Turku, Turku, Finland*

[6]*Aalto Design Factory, P.O. BOX 17700, FI-00076 Aalto, Finland*

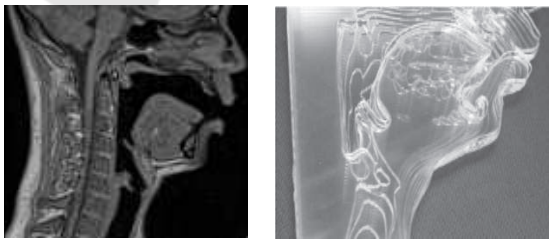
Keywords: MRI, 3D Image Processing, Automatic Surface Extraction, FEM Meshing, Physical Modelling.

Abstract: A procedure for the vectorisation and feature extraction of the human vocal tract is proposed. The raw data is obtained by high resolution 3D MRI. Because the amount of manual work in the data processing has been minimised, large datasets can be treated. The vectorised data can be used for both numerical as well as physical modelling of the vocal tract biophysics, including speech and applications in medicine.

1 INTRODUCTION

We present techniques, algorithms, and results for automatic vectorisation and feature extraction of grayscale voxel image files, produced by Magnetic Resonance Imaging (MRI). Our particular interest is in the extraction of the human vocal tract (VT) geometry. Numerous test subjects and large datasets are often involved, and manual data processing must therefore be minimised. This is the main motivation for developing custom software for VT extraction. Applications of anatomically accurate VT models range from computational modelling (see (Aalto et al., 2011; Dedouch et al., 2002; Hannukainen et al., 2007; Lu et al., 1993; Lacis, 2012) and the references therein) to fast prototyping of physical models for measurements that are impossible to carry out non-invasively using test subjects (see (Hirtum et al., 2011; Horáček et al., 2011; Šidlof et al., 2012; Takemoto et al., 2010)).

There is wide literature in image processing and feature extraction methods that have been applied in medical imaging; see, e.g., (Gonzalez and Woods, 2001; Criminisi et al., 2011). Several software solutions for processing of medical images exist such as the Vascular Modelling Toolkit (Vascular Modeling Toolkit, 2012) and MIMICS (Materialise, 2012) which was used in (Takemoto et al., 2010) for processing a small dataset of VT geometries. However, practical applications require specifically refined custom procedures that scale with the increasing size of the dataset.



(a) (b)

Figure 1: (a) A mid-sagittal section of a male subject while pronouncing vowel [œ]. (b) A plastic printout of the same vowel geometry, consisting of 24 sagittal planes of which 14 is shown here.

Compared to cardiovascular modelling and modelling of most internal organs, the extraction of the VT geometry has complications of its own. Teeth are missing from raw MRI data because osseous structures cannot be isolated from air volumes due to their low hydrogen content. This results in holes and other artefacts in vectorised images which must be corrected. As discussed below, the highly variable relative position of the mandible and the maxilla in different VT configurations must be taken into account

Aalto D., Helle J., Huhtala A., Kivelä A., Malinen J., Saunavaara J. and Ronkka T..
Algorithmic Surface Extraction from MRI Data - Modelling the Human Vocal Tract.
DOI: 10.5220/0004233302570260
In *Proceedings of the International Conference on Biomedical Electronics and Devices* (BIODEVICES-2013), pages 257-260
ISBN: 978-989-8565-34-1

in automated artefact removal. Separate imaging of teeth and their alignment with the soft tissue geometries from MRI is a difficult problem which is not discussed in this article.

2 SURFACE MODELS

In traditional medical imaging, it is sufficient to produce visualisations for inspection by trained radiologists. Numerical modelling methods require discrete representations such as tetrahedral meshes for FEM. We generate the mesh from a *solid triangular representation* of the desired surface which requires more data processing than surface extraction for visualization purposes only.

We propose a novel algorithm and software for surface and feature extraction of the 3D VT air volume Ω. The algorithm comprises the following steps:

1. **Pre-processing:** The voxel data is smoothed to remove noise.
2. **Initial Surface Extraction:** is carried out by extracting an *isosurface* corresponding to approximated air-tissue interface. An isosurface is a level set corresponding to a constant gray value, and it consists of triangular elements.
3. **Producing the Artefact Prior Model:** "Undesired artefacts", such as vertebrae and maxillae, are manually identified from the *initial surface* to produce two *prior models of the artefacts*, one for each maxilla.
4. **Removing Artefacts:** Artefact models are aligned with the initial surface using algorithms provided by PCL (Rusu and Cousins, 2011). Based on the location information thus obtained, the undesired artefacts are masked from the original voxel data.
5. **Final Surface Extraction:** An improved isosurface Γ is extracted from the artefact free, masked voxel data.
6. **Locating Boundaries:** The glottis, the velar port, and the mouth opening positions are located from Γ. These openings are covered, and they can be joined with Γ to produce a triangulated model for the full boundary $\partial\Omega$.

The raw data consists of MRI sequences that are stored as a DICOM files that comprising of 44 sagittal plane images: each image contains 128×128 pixels that are of size $d = 1.8$mm[1]. These planar images are aligned and stacked using the location data produced by the MRI machine. The stacked images form a 3D matrix of voxel data that represents the VT through grayscale values. The measurement setup for obtaining this data has been documented in (Aalto et al., 2011, Section 2.3).

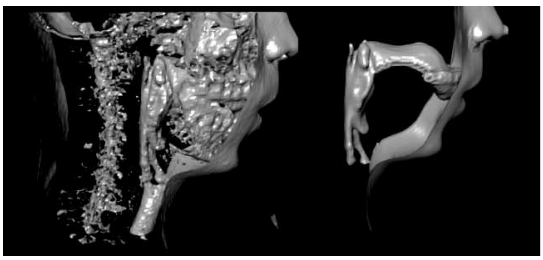


Figure 2: Shaded visualisation of the initial (left) and the final isosurfaces. The face is not part of the final computational geometry boundary $\partial\Omega$.

The isosurface extraction requires a threshold gray value that is determined from 2D bitmaps by estimating the VT boundary intersection heuristically. Since the air-tissue interface has a steep gray value gradient in the MRI voxel data, the extracted surface is not sensitive to small variations in the threshold value.

2.0.1 Artefact Modelling and Mesh Quality

The initial surface has a lot of artefacts as can be seen in Fig. 2. The alignment of the artefact models is carried out using the Point Cloud Library (PCL) (Rusu and Cousins, 2011). The information provided by PCL gives, in particular, the relative positions of the mandible and the maxilla which amounts to most of the positional variation between different VT configurations of the same test subject.

The production of the prior artefact models remains the most laborious piece of manual work where tools such as MeshLab (MeshLab, 2012) or Blender (Blender, 2012) can used. It takes about one hour to model the artefacts for one test subject but then the same artefact model can be used for all VT configurations, albeit only from the same test subject.

As the final result, an artefact free surface model of the VT geometry is obtained. One such geometry is shown in Fig. 2. The side lengths of the surface triangles are bounded above by $d\sqrt{3}$ where d is the voxel resolution of the MRI data. As reported in (Aalto et al., 2012, Section 4), the geometric error of the surface mesh is of order 0.5mm except in those parts of the model where MRI transparent teeth or the possibly open velar port cause crude error.

[1]The pixel number is always the same when using this sequence but the pixel size varies according to the physical dimension of the test subject.

3 CENTRELINE EXTRACTION

Many simplified computational physics models in tubular anatomic regions (such as flow mechanics in vascular structures and the acoustics of the VT) refer to centrelines and intersectional areas instead of the full 3D geometry. Our interest in centrelines and area functions of the VT stems from the generalised Webster's horn model (see, e.g., (Lukkari and Malinen, 2011), (Story et al., 1996)) which is a low frequency 1D approximation of the 3D wave equation on a tubular domain. For a discussion on hemodynamics models, see (Antiga, 2003) and the references therein.

Since centrelines are not uniquely defined on geometric grounds except in very simple cases, we first produce a candidate centreline for Ω. However, the "true" centreline depends on the application and the model of interest in a possibly intractable way. So as to Webster's model, we improve the model accuracy by choosing optimally from a family of centrelines that are normal to a fixed set of planar intersections as can be seen in Fig. 3.

Voronoi diagrams are used for the centreline extraction in Vascular Modelling Toolkit. Our approach is based on numerically solving the steady state heat equation with unit source $\Delta u = 1$ in Ω where $u = 0$ on VT walls, and $\frac{\partial u}{\partial \nu} = 0$ at mouth and glottis. In our setting, this can be done without much extra effort since FEM discretisation of Ω has already been produced for modelling VT acoustics, and a similar approach has been studied in (Vesom et al., 2008). The solution of the steady state heat equation can be obtained in linear time respect to the number n of nodes in the discretisation. The candidate centreline is, by definition, the ridge in the solution u. Finding the ridge is a pathfinding problem that can be solved in $O(n \log n)$ time.

4 PHYSICAL MODELS

We have produced pilot plastic models of the VT in 1:1 scale. We use these models for acoustic and flow measurements in order to augment and validate the numerical results from mathematical models.

We have experimented with hard plastic printouts only. Trials were performed using a 3D printer 3DTouch by Bits from Bytes, Ltd. The printing time for the full VT in 1:1 scale in PLA plastic was in excess of 12 hours when using a layer height of 0.25mm. The long printing time combined with the tendency of PVA to clog the extruder nozzle resulted in a disappointing print quality and a success rate of less than 25%. Stereolithography as in (Takemoto et al., 2010) or selective laser sintering is expected to produce higher quality models with better success rate than fused deposition methods.

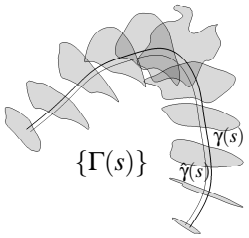


Figure 3: A subset of $\{\Gamma(s)\}$ and two centrelines $\gamma(s)$ and $\hat{\gamma}(s)$ corresponding to the chosen set of area functions.

The prototype model shown in Fig. 1(b) was produced by cutting sagittal intersection contours from 3 mm thick acrylic plate. We used Legend 36ext by Epilog Laser (http://www.epiloglaser.com) which is a CO_2 laser cutter with $P = 60W$ and two degrees of freedom. The cutting of all 24 sheets took under 2 hours.

The cutting angle of the sheets is always 90° in the model of Fig. 1(b), and this results in significant "stepping" of the VT boundary surface. The stepping can be reduced either by using thinner acrylic plate or, preferably, by varying the cutting angle so as to make the adjacent sheets fit to each other without steps. The variable cutting angle requires a CNC mill or a more advanced cutter.

Optically transparent models for Particle Image Velocimetry studies as in (Horáček et al., 2011) cannot be created simply by stacking transparent acrylic sheets but moulds for casting such models can.

5 CONCLUSIONS

We have developed algorithms for automatic surface detection and vectorisation of MR images of the head and neck area, especially concentrating on the vocal tract. A very large number of MR images are required for numerical model validation as well as medical applications. Hence, the amount of manual work must be minimised without sacrificing the data quality. Apart from some tasks related to manual artefact detection (see Step 3 in Section 2) and remaining open problems with efficient teeth modelling, the data processing can be carried out by a fully computerised procedure described in this work.

In addition to modelling purposes, physical printouts of vocal tract geometries may have future applications in reconstructive surgery and tissue engineering. Tissue grafts are produced by seeding and attachment of human cells into a *scaffold*. Scaffolds

must satisfy many material requirements due to biology (Sachlos and Czernuszka, 2003) as well as have the correct geometric shape, too.

ACKNOWLEDGEMENTS

The authors were supported by the Finnish Academy grant Lastu 135005, European Union grant Simple4All, Aalto Starting Grant, and Åbo Akademi Institute of Mathematics.

The current version of the software described in this paper can be obtained from the authors by request.

REFERENCES

Aalto, D., Aaltonen, O., Happonen, R.-P., Malinen, J., Palo, P., Parkkola, R., Saunavaara, J., and Vainio, M. (2011). Recording speech sound and articulation in MRI. In *Proceedings of BIODEVICES 2011*, Rome, Italy.

Aalto, D., Huhtala, A., Kivelä, A., Malinen, J., Palo, P., Saunavaara, J., and Vainio, M. (2012). How far are vowel formants from computed vocal tract resonances? arXiv:1208.5963, 13 pp.

Antiga, L. (2003). *Patient-Specific Modeling of Geometry and Blood Flow in Large Arteries*. PhD thesis, Politecnico di Milano.

Blender (2012). http://www.blender.org. Accessed Nov. 7th, 2012.

Criminisi, A., Shotton, J., and Konukoglu, E. (2011). Decision forests for classification, regression, density estimation, manifold learning and semi-supervised learning. Technical Report MSR-TR-2011-114, Microsoft Research.

Dedouch, K., Horáček, J., Vampola, T., and Černý, L. (2002). Finite element modelling of a male vocal tract with consideration of cleft palate. In *Forum Acusticum*, Sevilla, Spain.

Gonzalez, R. C. and Woods, R. E. (2001). *Digital Image Processing, 2nd Ed*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA.

Hannukainen, A., Lukkari, T., Malinen, J., and Palo, P. (2007). Vowel formants from the wave equation. *J. Acoust. Soc. Am. Express Letters*, 122(1):EL1–EL7.

Hirtum, A. V., Pelorson, X., and Estienne, O. (2011). Experimental validation of flow models for a rigid vocal tract replica. *J. Acoust. Soc. Am.*, 130(4):2128–2138.

Horáček, J., Uruba, V., Radolf, V., Veselý, J., and Bula, V. (2011). Airflow visualization in a model of human glottis near the self-oscillating vocal folds model. *Applied and Computational Mechanics*, 5:21–28.

Lacis, U. (2012). Modelling air flow in larynx. Master's thesis, Umeå University.

Lu, C., Nakai, T., and Suzuki, H. (1993). Finite element simulation of sound transmission in vocal tract. *J. Acoust. Soc. Jpn. (E)*, 92:2577 – 2585.

Lukkari, T. and Malinen, J. (2011). Webster's equation with curvature and dissipation. arXiv:1204.4075, 22 pp. + 5 pp. appendix.

Materialise (2012). Mimics. http://biomedical. materialise.com/mimics. Accessed Nov. 7th, 2012.

MeshLab (2012). Visual Computing Lab ISTI - CNR. http://meshlab.sourceforge.net/. Accessed Nov. 7th, 2012.

Rusu, R. B. and Cousins, S. (2011). 3D is here: Point Cloud Library (PCL). In *IEEE International Conference on Robotics and Automation (ICRA)*, Shanghai, China.

Sachlos, E. and Czernuszka, J. T. (2003). Making tissue engineering scaffolds work. Review: the application of solid freeform fabrication technology to the production of tissue engineering scaffolds. *Eur Cell Mater*, 5:29–39; discussion 39–40.

Story, B., Titze, I., and Hoffman, E. (1996). Vocal area functions from magnetic resonance imaging. *J. Acoust. Soc. Am.*, 100(1):537–554.

Takemoto, H., Mokhtari, P., and Kitamura, T. (2010). Acoustic analysis of the vocal tract during vowel productions by finite-difference time-domain method. *J. Acoust. Soc. Am.*, 128(6):3724–3738.

Vascular Modeling Toolkit (2012). http://www.vmtk.org. Accessed Nov. 7th, 2012.

Vesom, G., Cahill, N. D., Gorelick, L., and Noble, J. A. (2008). Characterization of anatomical shape based on random walk hitting times. In *In Proceedings of Mathematical Foundations of Computational Anatomy (MFCA 2008)*, New York.

Šidlof, P., Horáček, J., and Řidký, V. (2012). Parallel CFD simulation of flow in a 3D model of vibrating human vocal folds. *Computers and Fluids*. In press.