

 Open access • Journal Article • DOI:10.1080/10255840903505139

Finite element speaker-specific face model generation for the study of speech production — [Source link](#)

Marek Bucki, Mohammad Ali Nazari, Yohan Payan

Institutions: Centre national de la recherche scientifique, University of British Columbia

Published on: 15 Jul 2010 - Computer Methods in Biomechanics and Biomedical Engineering (Taylor & Francis Group)

Topics: Mesh generation

Related papers:

- [A biomechanical model of cardinal vowel production: muscle activations and the impact of gravity on tongue positioning.](#)
- [Coupled hard—soft tissue simulation with contact and constraints applied to jaw—tongue—hyoid dynamics](#)
- [Simulation of dynamic orofacial movements using a constitutive law varying with muscle activation](#)
- [A 3D model of muscle reveals the causes of nonuniform strains in the biceps brachii](#)
- [ArtiSynth: A Fast Interactive Biomechanical Modeling Toolkit Combining Multibody and Finite Element Simulation](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/finite-element-speaker-specific-face-model-generation-for-32e1jt3dv6>



HAL
open science

Finite element speaker-specific face model generation for the study of speech production

Marek Bucki, Mohammad Ali Nazari, Yohan Payan

► To cite this version:

Marek Bucki, Mohammad Ali Nazari, Yohan Payan. Finite element speaker-specific face model generation for the study of speech production. *Computer Methods in Biomechanics and Biomedical Engineering*, Taylor & Francis, 2010, 13 (4), pp.459-467. 10.1080/10255840903505139 . hal-00539097

HAL Id: hal-00539097

<https://hal.archives-ouvertes.fr/hal-00539097>

Submitted on 25 Nov 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Finite Element Speaker-Specific Face Model Generation for the Study of Speech Production

Marek Bucki¹, Mohammad Ali Nazari², and Yohan Payan^{1,3}

¹ TIMC-IMAG Laboratory, UMR CNRS 5525, University Joseph Fourier,
38706 La Tronche, France, Marek.Bucki@imag.fr

² ICP/GIPSA Laboratory, UMR CNRS 5216, 38000 Grenoble, France

³ PIMS-Europe, UMI CNRS 3069, University of British Columbia, Vancouver,
Canada

Abstract. In situations where automatic mesh generation is unsuitable, the Finite Element (FE) mesh registration technique known as Mesh-Match-and-Repair (MMRep) is an interesting option to quickly create a subject-specific FE model by fitting a predefined template mesh onto the target organ. The irregular or poor quality elements produced by the elastic deformation are corrected by a “mesh reparation” procedure ensuring that the desired regularity and quality standards are met. Here, we further extend the MMRep capabilities and demonstrate the possibility of taking into account additional relevant anatomical features. We illustrate this approach with an example of speaker’s face biomechanical model generation comprising face muscles insertions. While taking advantage of the *a priori* knowledge about tissues conveyed by the template model, this novel fast and automatic mesh registration technique makes it possible to achieve greater modeling realism by accurately representing the organ surface as well as inner anatomical or functional structures of interest.

1 Introduction

The physical speech signal is the consequence of the interaction between the mechanical structure of the articulatory apparatus (jaw, tongue, face, velum, vocal folds) and the forces that act on it. Part of these forces is the result of a central control while the other part depends on interactions between articulators and on peripheral feedback. Thus, to study the control of speech production from the articulatory or acoustic signals, some researchers have proposed to model separately the biomechanics of the articulators (with the anatomical, mechanical and dynamical properties of their tissues) and the neurophysiology of the control, [Laboissière et al.(1996)], [Perrier et al.(2003)]. Some clues were then provided in terms of biomechanical models of the jaw ([Laboissière et al.(1996)], [Hannam et al.(2008)]), tongue ([Payan and Perrier(1997),Dang and Honda(2004)], [Sanguineti et al.(1997)], [Perrier et al.(2003)], [Buchaillard et al.(2009)]), velum ([Payan et al.(2003)]) or vocal folds ([Hunter et al.(2004)], [Titze and Hunter(2007)]).

As regards the face, excepting the pioneer works in computer graphics for facial animation, [Lee et al.(1995)], [Sifakis et al.(2005)], and computer aided surgery, [Chabanas et al.(2003)], [Gladilin et al.(2004)], few works were provided, to our knowledge, in terms of accurate biomechanical modeling of the facial soft tissues for speech production study, i.e. with a focus on facial muscles activations and their control for a specific speech task. Some modeling works were provided with a focus on lips modeling, [Gomi et al.(2006)], [Kim and Gomi(2007)], while [Nazari et al.(2008)] started to model facial tissues deformations due to muscles activations for speech gestures.

Once models such as the one manually developed by [Nazari et al.(2008)] on a generic single example are able to simulate speech production, they will have to be geometrically adapted to morphologies of various speakers in order to compare measured data (speakers face deformations and acoustic signals) with numerical simulations provided by the models. This paper addresses such a perspective.

In terms of modeling, the mesh adaptation step requires a geometrical fit of the generic model onto each speaker anatomy. The generic face model proposed by [Nazari et al.(2008)] was built from data (CT exam) collected on a maxillo-facial patient. A three layers Finite Element (FE) mesh, symmetrical with respect to the sagittal plane, was manually designed to fit the patient skin (external mesh surface) and skull (internal mesh surface) geometries reconstructed from CT images. Inside the 3D mesh, Nazari and colleagues have manually defined a set of muscular macrofibres which are implemented in reference to anatomical key points, independently of the mesh structure, by cable (tension only) elements, as shown in Fig. 1.

Our goal is to provide a robust automatic procedure that will match the Nazari et al. generic FE mesh onto data collected on given speakers, while fulfilling the following constraints:

1. Each generated mesh must be as close as possible to the speaker's anatomy reconstructed from CT and/or MRI data; to this end, high accuracy must be guaranteed between original data and mesh geometry, with specific focus on the lip region which plays a crucial role in acoustical outcome;
2. The topology of the mesh has to be maintained so that the three anatomical layers are still represented (epidermis, dermis and hypodermis, from the external topmost layer to the internal one respectively);
3. The symmetry of the mesh along the sagittal plane has to be maintained;
4. The location of the speaker muscular structures must be inferred from the registration, assuming that some clues can be manually provided such as muscles insertion points onto the skull.

Our group has already provided partial responses to points 1) and 2) by proposing the Mesh-Match-and-Repair (MMRep) algorithm. This procedure automatically computes an elastic 3D transform that matches a generic mesh onto the anatomy of a new subject, while maintaining the topology and the FE regularity of the mesh. It was developed in the context of computer aided surgery and

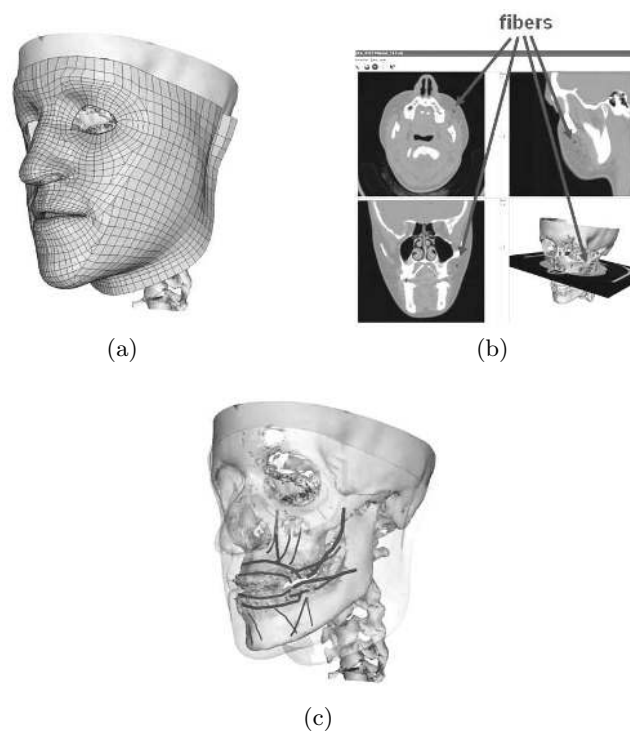


Fig. 1. (a) Generic FE mesh of the face soft tissues. (b) Interactive segmentation of muscles fibers (arrows) on CT data. (c) Location of some muscles involved in facial mimics (represented for the left side of the skull only).

evaluated on data acquired on maxillo-facial patients. This algorithm therefore did not focus specifically on the lips region.

Section 2 describes the MMRep algorithm. Section 3 explains how we suggest to extend the MMRep method and proposes a “MMRep-for-Speech” algorithm compatible with the above mentioned constraints 3) and 4). Finally, Section 4 demonstrates our new technique on image data collected from a speaker.

2 The Mesh-Match-and-Repair Algorithm

Finite Element mesh generation remains an important issue for subject specific biomechanical modeling. While some techniques, such as the widely used software Tetgen⁴, make automatic mesh generation possible, in most cases, manual mesh generation is preferred for better control over the elements layout that it provides. Yet, this option is time consuming and requires interaction with an expert operator.

The Mesh-Match-and-Repair (MMRep) algorithm discussed in this paper is a twofold procedure. First, the subject data and the generic (the term “Atlas” is often used) mesh surface nodes are registered using the **elastic deformation** procedure described by [Bucki(2008)]. This deformation is then applied to the inner Atlas nodes yielding a FE mesh that represents the modeled domain with sufficient accuracy.

As a consequence of this deformation, the Atlas elements may suffer distortions and become either “irregular elements” which make FE analysis impossible, or “poor quality elements” in which case the computation, although feasible, can exhibit numerical instabilities. To recover mesh **regularity** and reach an acceptable **quality** level a fast and robust **mesh reparation** procedure has been proposed by [Lobos(2009)]. The goal of the mesh reparation procedure is to optimize mesh regularity and quality criteria formulated as relaxation energy terms.

The MMRep procedure has been tested on 10 clinical cases in the field of orthopedics and 50 cases in the context of orthognathic surgery where it proved capable of automatically generating patient specific hexahedral dominant quality meshes within a couple of minutes time, [Bucki et al.(2009)]. In this study, the patients’ skin and skull surfaces were fully recovered from manually processed and segmented Computed Tomography (CT) scans, [Tilotta(2008)], and the face meshes registered so as to accurately represent both anatomical surfaces, as shown in Fig. 2.

3 The MMRep-for-Speech Algorithm

The MMRep-for-Speech algorithm proposed in this article provides the following functions:

⁴ <http://tetgen.berlios.de>

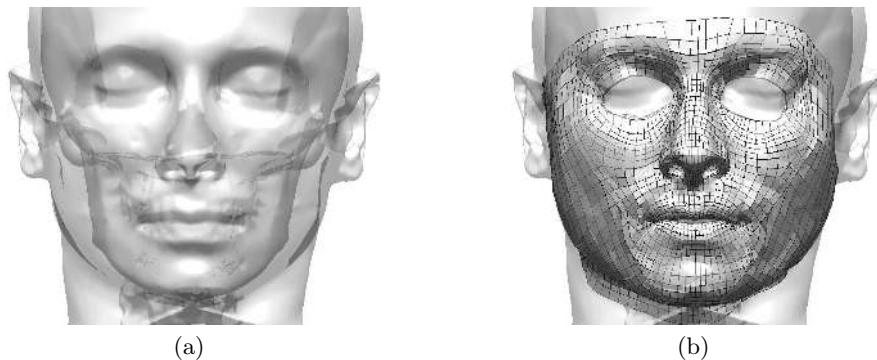


Fig. 2. MMRRep face registration. (a) Segmented skull and skin surfaces from a sample CT volume. (b) Registered FE mesh.

1. In order to improve the feature correspondence computed by the elastic registration, the ability to constrain the deformation with a set of control points which positions are known both in Atlas and speaker space, has been added to the mesh registration procedure (§3.1);
2. The elastic registration procedure which initially produced asymmetrical face meshes, as it can be seen on Fig. 2-b, has been further adapted to suit the prerequisites of speech production modeling by enforcing a symmetrical deformation field with reference to the central sagittal subject plane, while maintaining the regularity which makes it suitable for our mesh deformation procedure (§3.2);
3. As speech production simulation requires guaranteed mesh registration accuracy in specific anatomical areas such as the speaker’s lips region, a local registration refinement procedure has been implemented which strongly reduces the residual registration error (§3.3);
4. The face muscles that were integrated into the generic Atlas mesh proposed by [Nazari et al.(2008)] were automatically deformed by the computed elastic registration, therefore inferring the subject’s face muscles location (§3.4).

3.1 Control points driven registration

We define an elastic registration function as a mapping $R : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ that superimposes a set of Atlas mesh nodes A onto a speaker specific target data set S , which can either be a point cloud or a surface mesh. The computed elastic registration procedure complies with continuum mechanics conditions of motion, [Belytschko et al.(2006)], as R defines a C^1 -diffeomorphic, one-to-one and non-folding correspondence between geometric data sets, [Bucki(2008)], [Bucki et al.(2009)].

The iterative assembly of R is driven by the minimization of a “registration energy” E_{reg} which measures the similarity between the deformed Atlas point

set and the target data set S . As geometrical shape similarity is sought, E_{reg} is defined as the sum of Euclidean distances between deformed A and S :

$$E_{\text{reg}}(R) = \sum_{\mathbf{a} \in A} d(R(\mathbf{a}), S) \quad (1)$$

To increase registration performance, a distance map ([Saito and Toriwaki(1994)]) is computed based on the structure defined by S , and used for the evaluation of E_{reg} at each registration iteration.

In situations where the modeled organ geometry does not provide enough constraints for the registration to converge properly, for subjects with soft features, or where the face mesh is poorly reconstructed, the mesh deformation algorithm must be provided with additional data. This information can be supplied as a set of ‘‘control points’’ automatically or manually identified within both Atlas, $\bar{\mathbf{a}}_i, i \in [1, C]$, and subject’s space, $\bar{\mathbf{s}}_i, i \in [1, C]$. Any quantity of control points can be defined, but in our experience, a small number ($C = 5$) of key features seems sufficient.

To be effective, the influence of the small set of control points within the overall registration energy must be magnified by multiplicative ‘‘weight’’ factors $W_i, i \in [1, C]$, representing the strength of the attraction between the deformed control points $R(\bar{\mathbf{a}}_i)$ and their desired positions $\bar{\mathbf{s}}_i$ in subject space.

As the elastic deformation is being iteratively computed, the value of the weight factors for control points presenting inaccurate registration is increased until the desired accuracy standards are met. Within this new framework, the minimized registration energy defined in Eq. 1 is rewritten as:

$$E_{\text{reg}}(R) = \sum_{\mathbf{a} \in A} d(R(\mathbf{a}), S) + \sum_{i \in [1, C]} W_i d(R(\bar{\mathbf{a}}_i), \bar{\mathbf{s}}_i) \quad (2)$$

In order to provide greater control over features positioning during mesh registration, 5 skin control points were defined on half of Atlas and subject face prior to the symmetrical registration procedure described in §3.2. Fig. 3 displays an example of segmented half speaker’s face along with the 5 chosen control points defined in Atlas space (dark gray disks) and subject space (light gray disks). The dark dots represent the Atlas FE mesh nodes that need to be registered onto the subject’s skin and skull surfaces. The skull is not represented on the figure, for sake of clarity.

3.2 Symmetry enforcement

Speech production simulation run times can be greatly reduced by assuming sagittal symmetry and performing FE analysis on half of the face mesh. To this end, the registration must be constrained on the sagittal mid-plane so that the extension to the right half of the model of the deformation computed on the left half retains the continuity and smoothness properties mentioned above. This aspect differs from our previous application of the MMRep procedure in

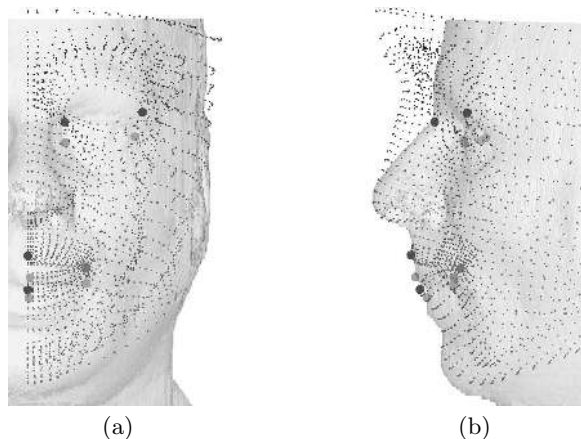


Fig. 3. Face registration control points in Atlas space (dark gray disks) and subject space (light gray disks): inner and outer corner of the left eye, left corner of the lips, upper and lower center of the lips. The dark dots represent the Atlas FE mesh nodes prior to registration. The light gray surface is the CT segmented subject’s skin, the reconstructed skull surface is not represented here. (a) Anterior view. (b) Sagittal view.

the context of orthognathic surgery where pathological patient morphology was essentially asymmetrical.

The first step in the symmetrical registration process is the proper identification of the pseudo-symmetry⁵ plane. The medical image volume sagittal plane is not necessarily a good candidate as the speaker may have tilted his or her head before the CT volume acquisition. A practical solution to this problem is proposed by [Chang and Park(2008)] and relies on rigid registration of the segmented skin surface onto itself after application of a rough planar symmetry (image sagittal mid-plane) and subsequent analysis of the correspondences between the surface points established by the rigid registration.

Once the sagittal pseudo-symmetry plane has been estimated, the elastic deformation R is computed as follows. The initial Atlas point set is embedded within a deformable “virtual elastic grid”. We arbitrarily set the shape of the grid to be the bounding box of the points, extended by a 10% margin. The considered deformation R is then formed by combinations of elementary deformations of the grid, $r_n, n \in [1, N]$, all having the desired regularity properties, much in the same way as proposed by [Rueckert et al.(2006)].

Each elementary deformation step n is performed by moving a single grid node and propagating the local deformation to the neighboring grid cells. The nodal displacement is transferred to the Atlas points located in the affected cells using a C^1 -diffeomorphic shape function $\varphi : \mathbb{R}^3 \rightarrow [0, 1]$ taking the value 1 at the displaced node and 0 at the boundary and outside the cells surrounding this node.

⁵ As subject data is not strictly speaking symmetrical.

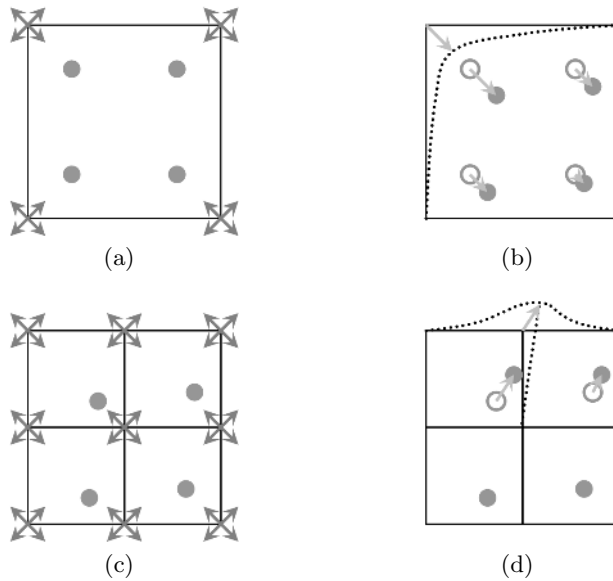


Fig. 4. Elastic registration overview. (a) A_0 at refinement level 1. (b) A_1 after deformation at level 1. (c) A_i at refinement level 2. (d) A_{i+1} after deformation at level 2.

The computation of each r_n is driven by the minimization of the registration energy defined in Eq. 2. As the registration progresses, the virtual grid is subdivided into smaller cells so that correspondence for finer anatomical details can be computed, [Bucki et al.(2009)]. Thus, the expression of the overall registration function is:

$$R = r_N \circ \dots \circ r_1$$

Fig. 4 illustrates in 2D our multi-grid iterative registration technique. The Atlas points A are the gray dots and S is not represented. In (a) the initial set $A_0 := A$ is embedded within the virtual grid (black square) discretized at the coarsest level 1 and the gradients of the energy, as defined in Eq. 2, are computed at the 4 mesh nodes. In (b) the nodal displacement yielding the highest energy decrease is applied to the grid and the embedded Atlas points move accordingly giving a new Atlas point set A_1 . This operation is repeated i times until no significant energy decrease can be generated at level 1. The grid mesh is then refined (c) and A_i is embedded within it. The energy gradients are computed at the 9 grid nodes, the best nodal displacement is applied (d) and the resulting Atlas point set A_{i+1} is computed. The registration iterations stop when a user defined highest grid refinement level is reached.

The enforcement of symmetrical deformation in this framework is quite straightforward. At each virtual grid refinement level, the grid nodes located on the face sagittal mid-plane are identified and their movements are constrained within the sagittal plane. As a consequence, the sagittal symmetry plane remains invariant for all considered elementary registration functions $r_n, n \in [1, N]$. Furthermore, thanks to the diffeomorphic nature of φ all continuity and smoothness properties are maintained throughout the sagittal mid-plane after the left half deformation has been transferred to the right side of the model.

3.3 Local registration refinement

The registration accuracy of the MMRep algorithm is strongly related to the resolution of the distance map used to speed-up the evaluations of E_{reg} at each step, which enables FE mesh generation within seconds. This resolution, limited by practical considerations such as distance map computation time and memory occupation, constrains the size of the smallest virtual grid cell reached during the top-down grid refinement process. Indeed, performing registration on a finer grid than the represented target set S is meaningless and leads to unnecessary local space distortions.

Yet, in the context of a speech production study, a submillimetric domain representation accuracy must be guaranteed in strategic areas such as the speaker’s lips region. To reduce the residual surface representation error in these regions after the initial deformation R has been applied, we propose to compute a finer local registration function Q based on more detailed lips geometry information. This local registration adjustment relies on a finer distance map and hence, enables smaller virtual grid cells to be used, while requiring acceptable computational overhead.

To maintain the consistency of the resulting elastic registration $Q \circ R$, continuity and smoothness must be preserved not only on the local sub-domain subject to refinement but also at the overall face scale. Again, this is accomplished by controlling the displacements of the virtual grid nodes at each grid refinement level.

Fig. 5 illustrates our local elastic registration refinement strategy. First, the overall registration R is computed over the raw virtual grid (black). Then, a finer distance map is used to increase the surface representation accuracy in the region surrounded by the finer grid (dashed) while the points located outside remain unchanged. The continuity and smoothness of $Q \circ R$ through the boundary between both grids is preserved by considering a registration function Q that leaves all its underlying grid boundary nodes fixed. The definition of the shape functions φ used to transfer the virtual grid nodes displacements to the enclosed Atlas mesh nodes ensures that the overall registration function $Q \circ R$ retains the smoothness properties discussed in §3.1.

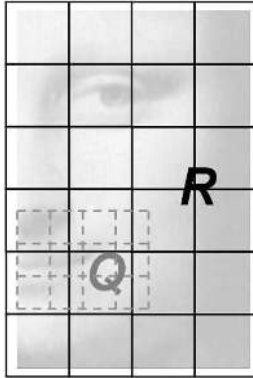


Fig. 5. Registration refinement. Black grid: overall half face registration R . Dashed grid: accurate lips registration Q . The boundary between the overall deformation domain and the lips sub-domain is fixed.

3.4 Speaker muscles location inferred by the elastic registration

Muscle courses are identified from CT data and anatomical facts. Muscle insertion points can be recognized from the landmarks on the skull surface. Muscles meet points can also be used to define muscle tracks, especially for those which do not have skull insertion points. Since muscles appear as fiber bundles, the course of these fibers can be modeled by a few linear piecewise continuous tracks.

The number of tracks per muscle depends on the extent of each muscle. For example in orbicularis oris two tracks are attributed for peripheralis and marginalis parts. Each track is characterized by a series of points which are picked up from the CT data based on anatomical facts and relying on surgeon's expertise, [Nazari et al.(2008)].

In our approach, the muscle tracks identified in the Atlas mesh by an expert operator are deformed by the elastic registration function computed as described above. The muscle tracks structure is preserved by the deformation while the inferred subject specific insertions points match the considered speaker's anatomy.

4 Results

In order to check its applicability in situations where less attention is paid to the regularity of the reconstructed subject anatomical structures, the MMR_{ep}-for-Speech algorithm has been tested on raw clinical data, as opposed to the manually processed skull and face meshes used in our previous study, [Bucki et al.(2009)]. Popular surface segmentation and reconstruction techniques often yield uneven surface meshes, possibly exhibiting segmentation artifacts. Automatic clean-up is not always satisfactory and a manual time consuming post-processing is required in most cases. As it is usually difficult to achieve smooth surface reconstruction, FE mesh registration procedures must be able to deal with this limitation.

The new algorithm was tested on a sample speaker data set. Bone and skin surfaces were automatically extracted from the CT volume using the Marching Cubes method, [Lorensen and Cline(1987)]. The segmented and reconstructed skull surface comprises 263,994 vertices and 421,410 triangles, while the skin model has 128,011 vertices and 253,654 triangles. Both skin and skull surfaces are taken into account in the subject specific FE mesh produced by the MMRep algorithm, as illustrated in Fig. 6.

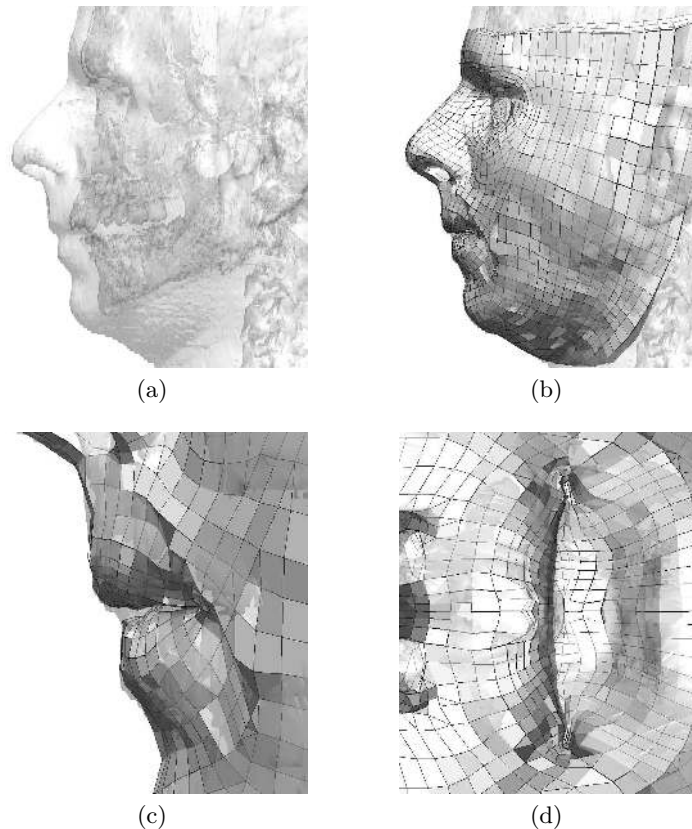


Fig. 6. Sample speaker data. (a) Bone and skin surfaces reconstructed from CT. (b) FE mesh produced by the MMRep-for-Speech procedure. (c) Sagittal close-up view on the lips region. (d) Anterior close-up view on the lips region.

After overall mesh registration R and subsequent reparation, the surface representation errors measured as Atlas mesh nodes-to-surface distances are given in table 1. Mesh registration and reparation time was less than one minute. The registration was computed using a point-to-point distance map with a resolution of 1.5mm but the errors reported below are measured using the actual triangular

surface meshes. An overall submillimetric accuracy was achieved and the maximal errors are mainly due, for the skull, to an approximative representation of the temporal bones in the Atlas, and, for the skin, to the absence of ears in the generic mesh, as can be seen in Fig. 6-b.

| Surface | Mean | Max. | σ |
|---------|------|------|----------|
| Skull | 0.5 | 12.4 | 0.5 |
| Skin | 0.3 | 19.2 | 0.4 |

Table 1. Mesh registration errors, in millimeters, measured as distances between the Atlas mesh skull or skin nodes and their corresponding anatomical surface (mean, maximal error and standard deviation).

As for the face control points registration accuracy, the errors measured on the set of 5 features presented in Fig. 3 are: mean err. = 0.7 mm, max. err. = 1.0 mm and $\sigma = 0.2$ mm. These figures could be improved by using a distance map with a greater resolution or stronger control points weights. However this seems unnecessary as the purpose of control points is merely to ensure that proper features geometry is reconstructed by the registration process. Although the Atlas control points near the lips do not perfectly overlap their subject specific counterparts, it is clear that the lips geometry has been perfectly captured by the registration procedure, as Fig. 6-c and d show.

Within the Atlas mesh, a subset of 350 skin nodes has been marked as “lips nodes”. These nodes were used, first, to measure the mesh registration accuracy in this region and, second, to define a sub-region where registration refinement would take place, should the achieved accuracy be unsatisfactory.

Table 2 shows the lips registration accuracy after overall half face registration (i.e. without local registration refinement), after local registration refinement using a 0.5mm resolution distance map and finally after reparation of the locally registered mesh.

The lips surface representation accuracy is higher after local registration refinement, with a maximal error of 0.15 mm. Nevertheless, as the resulting mesh needs to undergo the reparation procedure, i.e. regularity recovery and quality optimization, some lips surface nodes positions are altered resulting in a slight accuracy loss, shown in the bottom line of table 2. Nevertheless, the resulting FE mesh presents an overall submillimetric mean accuracy and, in the particular lips region, a maximal surface representation error of 0.29 mm which is an acceptable accuracy standard for speech production modeling.

As stated initially, the Atlas registration procedure was able to recover an approximate localization of the speaker’s face muscles, as illustrated in Fig. 7-b. Unfortunately, in the absence of anatomical information about the position of the actual muscle insertions, no quantitative comparison could be made between the positions of these anatomical landmarks and those inferred by the elastic registration algorithm.

| Step | Mean | Max. | σ |
|----------|------|------|----------|
| Overall | 0.24 | 1.02 | 0.19 |
| Refined | 0.02 | 0.15 | 0.03 |
| Repaired | 0.02 | 0.29 | 0.04 |

Table 2. Lips registration errors, in millimeters, after overall half face registration, refined lips registration (before mesh reparation) and reparation of the refined mesh registration.

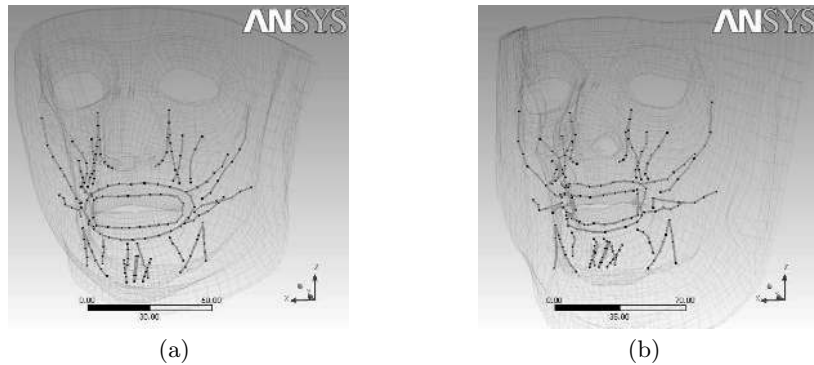


Fig. 7. (a) Muscles layout within the generic Atlas mesh, shown in transparency. (b) Muscles location inferred by the elastic deformation computed between the Atlas mesh and the considered subject's face.

Finally the produced mesh could be used to simulate the contraction of two muscles on the subject's face. The effects of the contractions of zygomaticus and orbicularis oris peripheralis are respectively shown in Figures 8 and 9. As can be observed, the activation of zygomaticus draws the corner of the mouth upward and laterally and activation of orbicularis oris peripheralis shows lips protrusion which qualitatively fits the anatomical facts even if the magnitude of protrusion seems too small.

5 Conclusions and Discussion

A series of extensions of MMRep, our fast and robust mesh generation technique, has been proposed and successfully tested on an illustrative case reproducing the constraints likely to be encountered in a speech production study.

The proposed algorithm offers an interesting balance between manual FE mesh generation with fully controlled element features and automatic "from scratch" mesh generation, by taking as starting point a user defined, tailor made generic mesh. Our procedure is capable of automatically generating within minutes a subject specific mesh directly exploitable by a commercial FE software such as ANSYS Workbench (ANSYS Inc., USA).

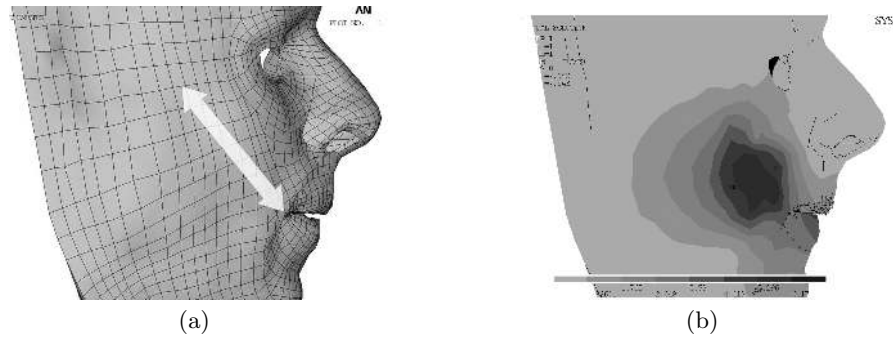


Fig. 8. Effects of zygomaticus muscle (white arrow) contraction on the face mesh. (a) Rest position. (b) Deformed face.

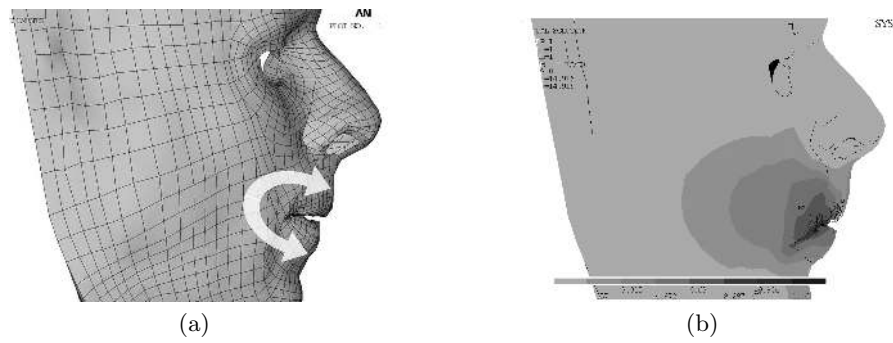


Fig. 9. Effects of orbicularis oris muscle (white arrow) contraction on the face mesh. (a) Rest position. (b) Deformed face.

Among the proposed extensions, only the definition of the control points in both Atlas and speaker space required user interaction. Yet the 5 control points used here could easily be identified by manual selection in less than a minute. If necessary, for these simple face features an automatic extraction algorithm may also be envisaged.

Future work includes enhancement of the mesh reparation procedure in order to take into account the surface representation constraint and avoid accuracy loss such as that shown in table 2. The current algorithm considers all nodes indistinctly and the search for an acceptable mesh configuration is performed by moving in priority the nodes presenting the worst regularity or quality conditions. A first enhancement would be to begin this search by moving the inner mesh nodes whose locations do not affect the surface representation accuracy. Furthermore, the displacements of surface nodes could also be considered, but should be constrained by the previously recovered subject surface or its vicinity.

The elastic registration algorithm used here preserves a certain regularity of space but can also reproduce finer details when the subject data are irregular, for morphological reasons or due to the presence of noise. This capability explains the jagged aspect of some muscles fibers in Fig. 7-b as opposed to the regular shapes featured in Fig. 7-a. Indeed, for the poorly reconstructed speaker data presented here, a point-to-point distance measure was used to evaluate the registration energy, mainly to reduce computation times. A point-to-surface measure would probably define a smoother energy landscape and contribute to preserving muscle regularity throughout the registration process.

The muscle deformation regularity could be further enhanced by considering during the registration computation a mechanical regularization term, as suggested by [Bucki et al.(2009)], which would specifically reflect the bending of the muscle fibers. As a consequence, the muscles network embedded within the face mesh would behave like a wire-frame structure preserving the surrounding tissues as well as its own smoothness.

Finally, from an anatomical point of view, a validation of the position of muscle insertions inferred from the registration procedure must be carried out based on actual muscle insertion coordinates in speaker space. If it appeared to be insufficiently accurate, then these insertion points could be used as registration control points, which should lead, as suggested above, to realistic muscle positioning within the reconstructed subject specific mesh and more accurate biomechanical modeling of the face's soft tissues.

6 Acknowledgments

The authors wish to thank Mr. Badin for his remarkable anatomy.

References

- [Belytschko et al.(2006)] Belytschko T, Liu WK, Moran B. 2006 Nonlinear Finite Elements for Continua and Structures. Wiley.

- [Buchillard et al.(2009)] Buchillard S, Perrier P, Payan Y. 2009. Muscle activations and the impact of gravity on tongue positioning during cardinal vowel production: a modeling study. Manuscript accepted for publication in *Journal of Acoustical Society of America* .
- [Bucki(2008)] Bucki M. 2008. Modélisation Biomécanique des Tissus Mous du Cerveau et Développement dun Neuronavigateur Permettant la Prise en Compte Per-Opératoire du Brain-Shift Université Joseph-Fourier - Grenoble I, France.
- [Bucki et al.(2009)] Bucki M, Lobos C, Payan Y. 2009. A Fast and Robust Patient Specific Finite Element Mesh Registration Technique: Application to 60 Clinical Cases. Manuscript submitted to *Medical Image Analysis* .
- [Chabanas et al.(2003)] Chabanas M, Luboz V, Payan Y. 2003. Patient specific finite element model of the face soft tissues for computer-assisted maxillofacial surgery. *Medical Image Analysis* 7:131–151.
- [Chang and Park(2008)] Chang M, Park SC. 2008. Reverse engineering of a symmetric object. *Computers and Industrial Engineering* 55:311–20.
- [Dang and Honda(2004)] Dang J, Honda K. 2004. Construction and control of a physiological articulatory model. *The Journal of the Acoustical Society of America* 115:853–870.
- [Gladilin et al.(2004)] Gladilin E, Ivanov A, Roginsky V. 2004. A framework for biomechanical simulation of cranio-maxillofacial surgery interventions In: *Proceedings of ISMS 2004* Cotin S Metaxas D. , editors. pp. 287–294.
- [Gomi et al.(2006)] Gomi H, Nozoe J, Dang J, Honda K. 2006. “A physiologically based model of perioral dynamics for various lip deformations in speech articulation,” *Speech Production : Models, Phonetic Processes, and Techniques*. Psychology Press. pp. 119–134.
- [Hannam et al.(2008)] Hannam AG, Stavness I, Lloyd JE, Fels S. 2008. A dynamic model of jaw and hyoid biomechanics during chewing. *Journal of Biomechanics* 41(5):1069–1076.
- [Hunter et al.(2004)] Hunter EJ, Titze IR, Alipour F. 2004. A three-dimensional model of vocal fold abduction/adduction. *Journal of Acoustical Society of America* 115(4):1747–1759.
- [Kim and Gomi(2007)] Kim K, Gomi H. 2007. Model-based investigation of control and dynamics in human articulatory motion. *Journal of system design and dynamics* 1:558–569.
- [Laboissière et al.(1996)] Laboissière R, Ostry DJ, Feldman AG. 1996. The control of multi-muscle systems: Human jaw and hyoid movements. *Biological Cybernetics* 74(3):373–384.
- [Lee et al.(1995)] Lee Y, Terzopoulos D, Waters K. 1995. Realistic modeling for facial animation In: *Proceedings of SIGGRAPH'95* Mair SG Cook R. , editors. New York. ACM Press. pp. 55–62.
- [Lobos(2009)] Lobos C. 2009. Amélioration des Techniques de Génération de maillages 3D des structures anatomiques humaines pour la Méthode des Eléments Finis Université Joseph-Fourier - Grenoble I, France.
- [Lorensen and Cline(1987)] Lorensen WE, Cline HE. 1987. Marching cubes: A high resolution 3D surface construction algorithm In: *Proceedings of the 14th annual conference on Computer graphics and interactive techniques - SIGGRAPH'87*. pp. 163–169.
- [Nazari et al.(2008)] Nazari MA, Payan Y, Perrier P, Chabanas M, Lobos C. 2008. A continuous biomechanical model of the face: a study of muscle coordinations for speech lip gestures In: *Proceedings of the 8th International Seminar on Speech Production - ISSP'08*. pp. 321–324.

- [Payan et al.(2003)] Payan Y, Pelorson X, Perrier P. 2003. Physical modelling of the airflow-walls interactions to understand the sleep apnea syndrome. *Lecture Notes in Computer Science* 2673:261–269.
- [Payan and Perrier(1997)] Payan Y, Perrier P. 1997. Synthesis of V-V sequences with a 2D biomechanical tongue model controlled by the Equilibrium Point Hypothesis. *Speech Communication* 22:185–205.
- [Perrier et al.(2003)] Perrier P, Payan Y, Zandipour M, Perkell J. 2003. Influences of tongue biomechanics on speech movements during the production of velar stop consonants: a modeling study. *Journal of Acoustical Society of America* 114(3):1582–1599.
- [Rueckert et al.(2006)] Rueckert D, Aljabar P, Heckemann RA, Hajnal JV, Hammers A. 2006. Diffeomorphic registration using B-splines. In: *Proceedings of Medical Image Computing and Computer-Assisted Intervention MICCAI'06*. pp. 702–9.
- [Saito and Toriwaki(1994)] Saito T, Toriwaki JI. 1994. New algorithms for Euclidean distance transformations of an n-dimensional digitized picture with applications. *Pattern Recognition* 27:1551–65.
- [Sanguineti et al.(1997)] Sanguineti V, Laboissière R, Payan Y. 1997. A control model of human tongue movements in speech. *Biological Cybernetics* 77(11):11–22.
- [Sifakis et al.(2005)] Sifakis E, Neverov I, Fedkiw R. 2005. Automatic Determination of Facial Muscle Activations from Sparse Motion Capture Marker Data. In: *ACM Transactions on Graphics (SIGGRAPH Proceedings), TOG 24*. pp. 417–425.
- [Tilotta(2008)] Tilotta F. 2008. Contribution à la reconstitution faciale en médecine légale : Proposition d'une nouvelle méthode statistique. Université Paris Sud 11, France.
- [Titze and Hunter(2007)] Titze I, Hunter E. 2007. A two-dimensional biomechanical model of vocal fold posturing. *Journal of Acoustical Society of America* 121(4):2254–2260.