LOME LABORATORY OF OPTICS AND EXPERIMENTAL MECHANICS

inegi

Universidade do Porto Faculdade de Engenharia FEUP

Segmentation of the Vocal Tract in Magnetic Resonance Images using Deformable Models

Maria J. M. Vasconcelos, Sandra M. Ventura,
Diamantino R. S. Freitas, **João Manuel R. S. Tavares**

Faculdade de Engenharia da Universidade do Porto, Portugal

maria.vasconcelos@fe.up.pt, smr@estsp.ipp.pt, dfreitas@fe.up.pt, **tavares@fe.up.pt**

6th World Congress on Biomechanics

in conjunction with

14th International Conference on Biomedical Engineering (ICBME) &

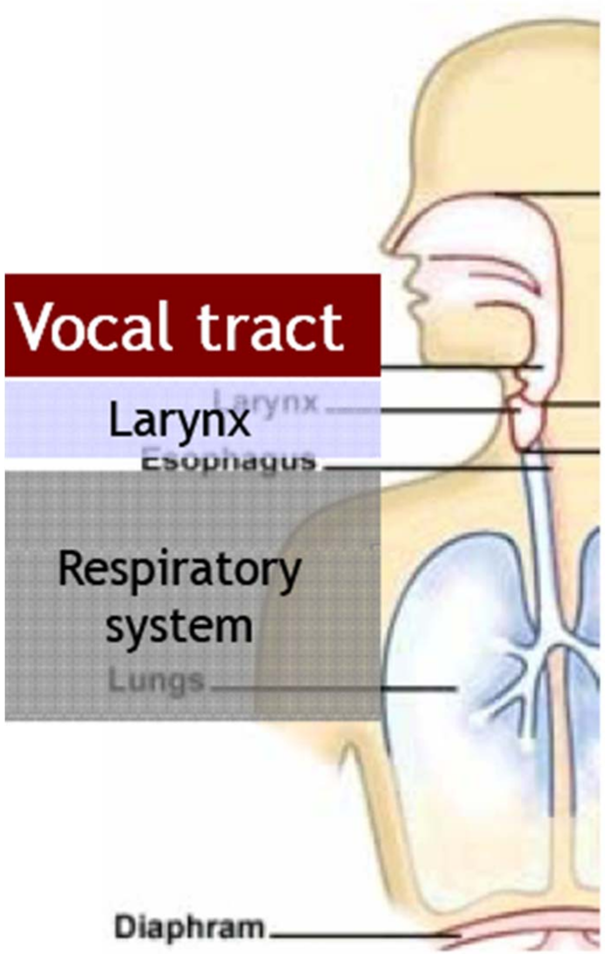
5th Asian Pacific Conference on Biomechanics (APBiomech)

Outline

- Introduction
  - Speech production, Vocal tract, Goals
- Data and Methods
  - MRI protocol, Speech corpus, Vocal tract's shape modelling (PDM - Point Distribution Model – Building process), Vocal tract's shape segmentation (ASM - Active Shape Model and AAM - Active Appearance Model)
- Results
  - Landmark points, Modes of variation, Segmentation
- Conclusions

Introduction: Speech production

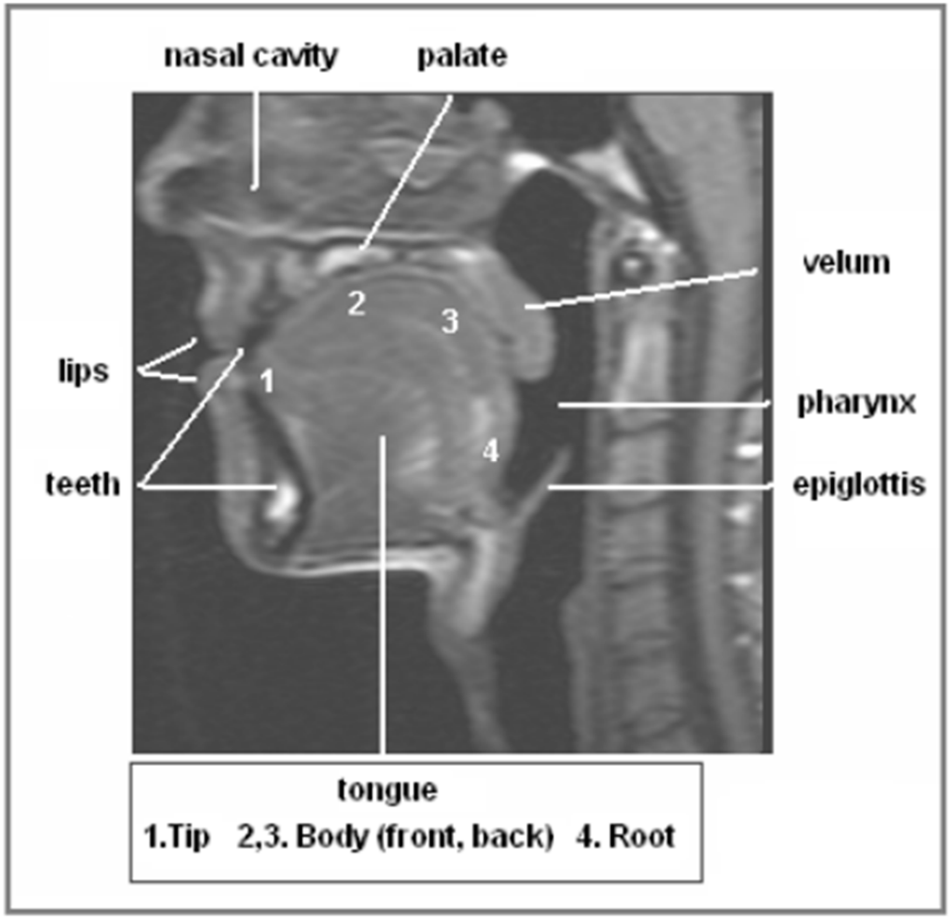
- Speech production involves complex and individual mechanisms
- It includes three main steps
  - Respiration (air supply – Activation)
  - Phonation (sound origin – Vibration)
  - Articulation (sound modulation – Articulation and Resonance)



Ventura et al. (2009), Application of MRI and Biomedical Engineering in Speech Production Study, Computer Methods in Biomechanics and Biomedical Engineering 12(6):671-681

Introduction: Vocal Tract

- Main shape features
  - Non-linear shape ("L-shaped tube")
  - Significant length
- Articulators' behavior
  - passive (i.e. palate, pharynx)
  - active (i.e. tongue, lips, velum)



Ventura et al. (2010), Towards Dynamic Magnetic Resonance Imaging of the Vocal Tract during Speech Production, Journal of Voice, doi:10.1016/j.jvoice.2010.01.014 (in press)

Introduction: Goals

- **Characterize** and **further reconstruct the vocal tract during** the **articulation** of (25) **European Portuguese sounds** by using techniques of image processing and analysis on magnetic resonance images
- **Automatically segment the vocal tract** in magnetic resonance images

Data and Methods

- Image acquisition
  - Siemens Magneton Symphony 1.5T system
  - Head array coil
- Speech corpus
  - One training subject (male) in supine position, without speech disorders
  - Twenty five European Portuguese sounds

Vasconcelos et al. (2010), Towards the Automatic Study of the Vocal Tract from Magnetic Resonance Images, Journal of Voice, doi:10.1016/j.jvoice.2010.05.002 (in press)

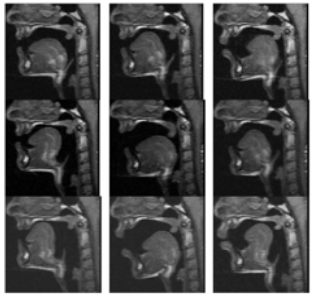
Data and Methods: MRI protocol

- Slices in sagittal orientation
- T1-weighted images using Turbo Spin Echo sequences
- Time acquisition: 10 sec. (compromise between image resolution and time for sustained articulation tolerable by the subject)
- Static study (sustained articulation during acquisition)
- One image was acquired per each sound (25)

Ventura et al. (2009), Application of MRI and Biomedical Engineering in Speech Production Study, Computer Methods in Biomechanics and Biomedical Engineering 12(6):671-681

Data and Methods: Point Distribution Model (Building process)

Training set (21 images)



(N) **Objects sampled** by a set of n (25) **landmark points**

Building the **vector of landmark points' co-ordinates**:

$$x_i = (x_{i1}, \ldots, x_{in}, y_{i1}, \ldots, y_{in})^T$$

with $i = 1 \ldots N$

Rigid registration of all objects in the training set (considering rotation, translation and scale)

Compute the mean shape ($\bar{x}$) **and the variation modes** (by applying a Principal Component Analysis on the landmark points' co-ordinates)

Pontual Distribution Model

$$x = \bar{x} + Pb$$

with P the matrix with the first t eigen-vectors, and b the vector of weights for each variation mode

Vasconcelos et al. (2008), Methods to Automatically Built Point Distribution Models for Objects like Hand Palms and Faces Represented in Images, Computer Modeling in Engineering & Sciences 36(3):213-241

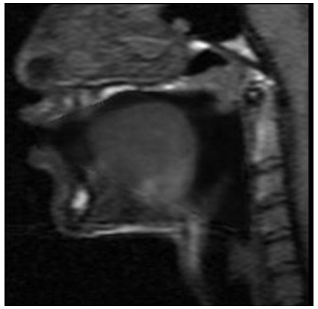
Data and Methods: Active Shape Model (Segmentation process)

Shape Model (Point Distibution Model)

+

Intensity profile of each landmark point

New image



Active Shape Model

Goal: match the model into a new image (**image segmentation**)

By considering iteratively the steps:
- Search for a better position for each landmark point (trough Mahalanobis distance)
- Update the model's parameters in order to fit the new landmark points' positions

Vasconcelos et al. (2008), Methods to Automatically Built Point Distribution Models for Objects like Hand Palms and Faces Represented in Images, Computer Modeling in Engineering & Sciences 36(3):213-241

Data and Methods: Active Appearance Model (Seg. process)

Shape Model

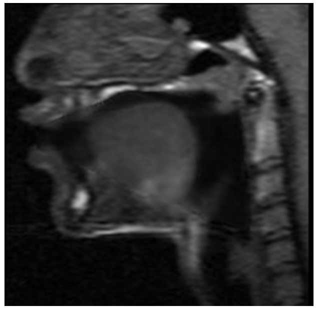
\+

Texture Model

Appearance Model

Combination of a shape model (geometric) with a texture model (i.e. intensity information)

New image



Active Appearance Model

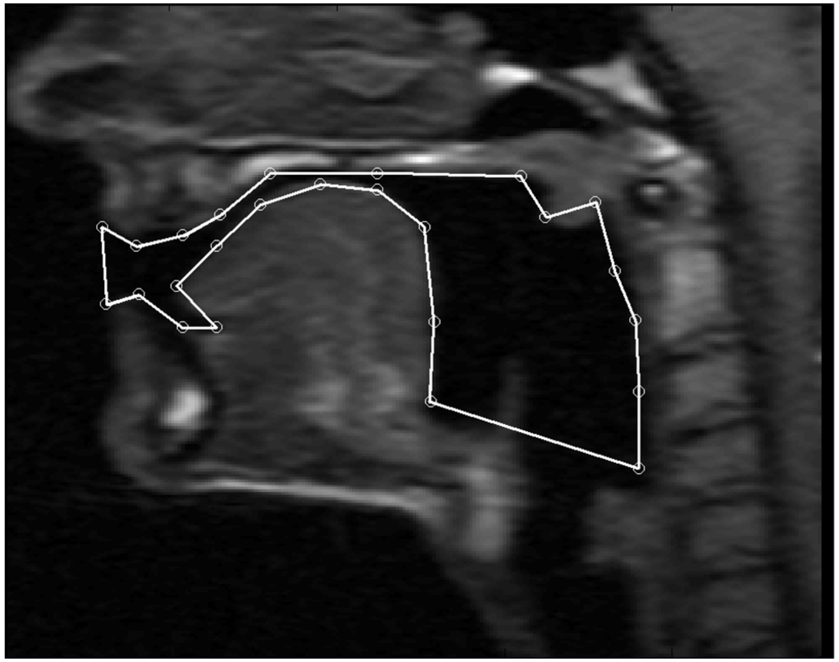
Goal: match the model into a new image (**image segmentation**)

By considering iteratively the steps:
- Compute the residual errors between the new shape and the model built
- Use the model to predict the changes on its current parameters in order to reduce the residual errors

Vasconcelos et al. (2008), Methods to Automatically Built Point Distribution Models for Objects like Hand Palms and Faces Represented in Images, Computer Modeling in Engineering & Sciences 36(3):213-241

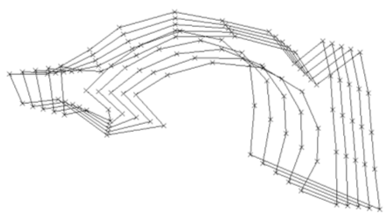
Results: Landmark points

- 25 landmark points were used
  - 4 points at the lips
  - 3 points corresponding to the lingual frenulum and tongue's tip
  - 7 points equally spaced along the surface of the tongue
  - 7 points along the surface of the hard palate, symmetric with tongue points
  - 1 point at the velum
  - 3 points equally spaced at the posterior margin of the oropharynx

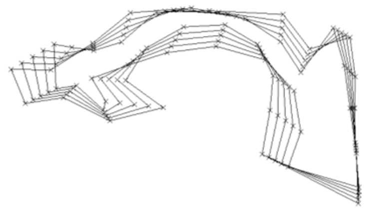


Vasconcelos et al. (2010), Towards the Automatic Study of the Vocal Tract from Magnetic Resonance Images, Journal of Voice, doi:10.1016/j.jvoice.2010.05.002 (in press)

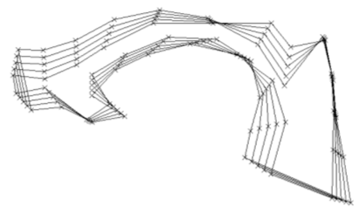
Results: Shape variation modes (-2σ, -σ, 0, +σ, +2σ)



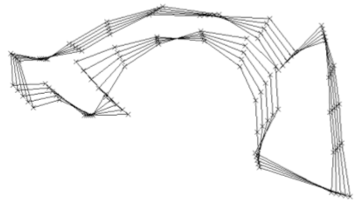
1st mode of variation



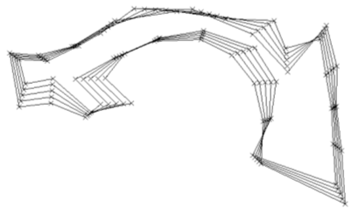
2nd mode of variation



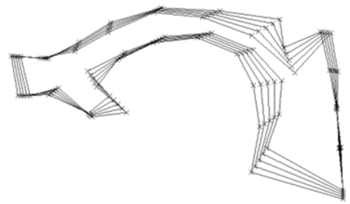
3rd mode of variation



4th mode of variation



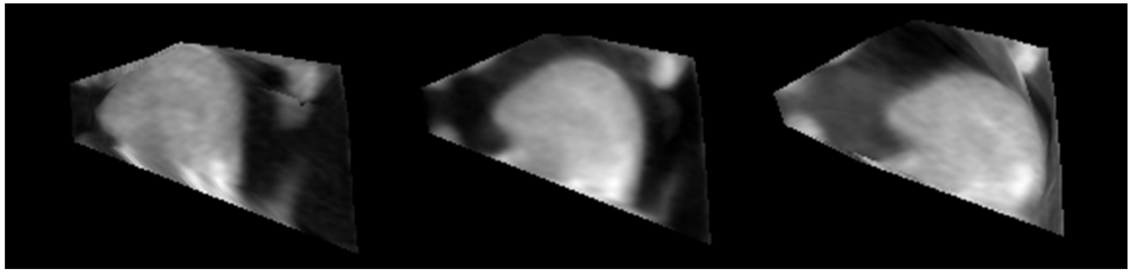
5th mode of variation



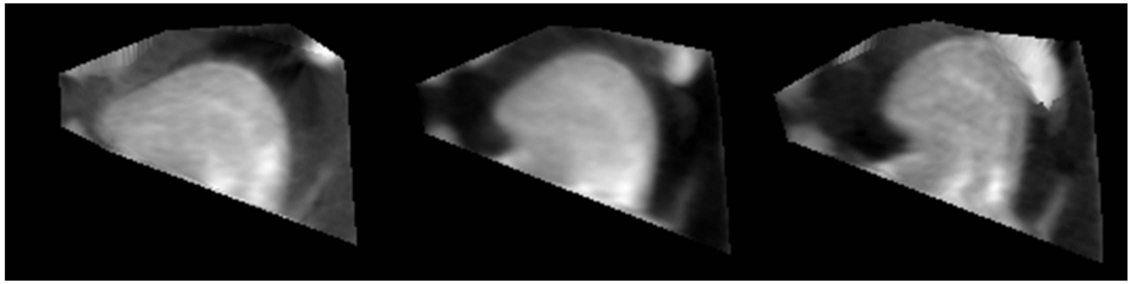
6th mode of variation

Vasconcelos et al. (2010), Using Statistical Deformable Models to Reconstruct Vocal Tract Shape from Magnetic Resonance Images, Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine, doi:110.1243/09544119JEIM767 (in press)

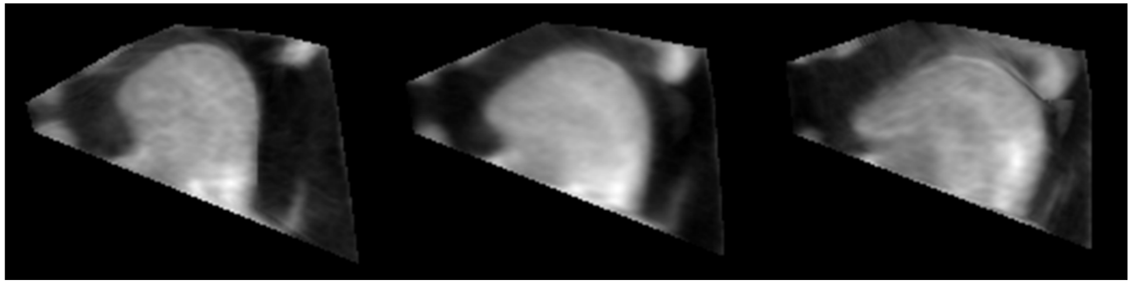
Results: Appearance variation modes (-σ, 0, +σ)



1st mode of variation



2nd mode of variation

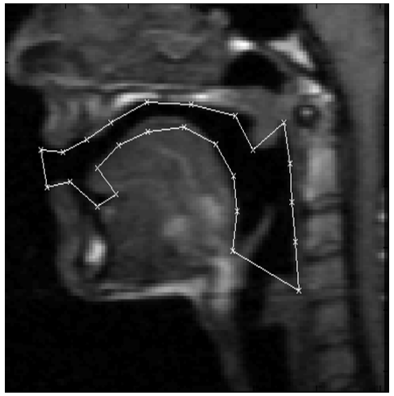


3rd mode of variation

Results: Vocal tract's shape segmentation

- Active Shape Model


Initial position

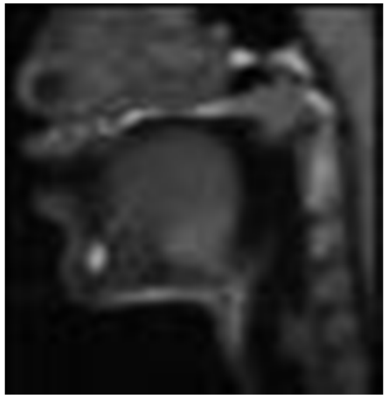

4th Iteration

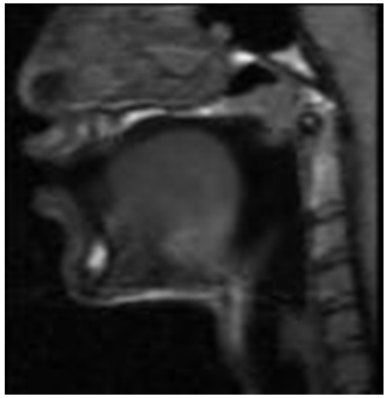

9th Iteration

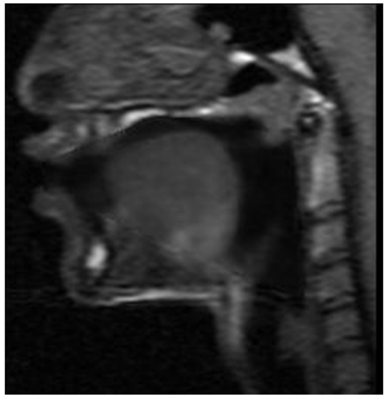

14th Iteration

- Active Appearance Model


1st Iteration


7th Iteration


12th Iteration


20th Iteration

Results: Vocal tract's shape segmentation

- Errors between automatic/manual segmentations (in pixel)

	Image 1	Image 2	Image 3	Image 4
Active Shape Model	9.99 ± 5.76	9.89 ± 4.43	11.54 ± 6.36	14.23 ± 7.66
Active Appearance Model	4.90 ± 2.42	10.21 ± 5.09	8.98 ± 4.80	9.91 ± 3.95

Vasconcelos et al. (2010), Towards the Automatic Study of the Vocal Tract from Magnetic Resonance Images, Journal of Voice, doi:10.1016/j.jvoice.2010.05.002 (in press)

Conclusions

- We applied statistical models in MR images to analyze the vocal tract's shape in the articulation of some European Portuguese Sounds and used the models built to segment the vocal tract's shape in new images
- From the results experimental results, we can conclude that the statistical models built can extract the main characteristics of the movements of vocal tract effectively
- Additionally, the models built revealed to be efficient to segment the vocal tract's shape from MR images
- Therefore, the models built can be accurate and efficient tools to be used towards the automatic study of the vocal tract from MR images in, for example, speech simulation or rehabilitation

Acknowledgements

- Authors would like to thank the support of the PhD grants SFRH/BD/28817/2006 and SFRH/PROTEC/49517/2009, from Fundação para a Ciência e Tecnologia (FCT), in Portugal.
- This work was partially done in the scope of the project "Methodologies to Analyze Organs from Complex Medical Images – Applications to Female Pelvic Cavity" - PTDC/EEA-CRO/103320/2008, financially supported by FCT.

LOME LABORATORY OF OPTICS AND EXPERIMENTAL MECHANICS

inegi

Universidade do Porto Faculdade de Engenharia FEUP

6th World Congress on Biomechanics
in conjunction with
14th International Conference on Biomedical Engineering (ICBME) &
5th Asian Pacific Conference on Biomechanics (APBiomech)

Thank you!

Segmentation of the Vocal Tract in Magnetic Resonance Images using Deformable Models

Maria J. M. Vasconcelos, Sandra M. Ventura,
Diamantino R. S. Freitas , João Manuel R. S. Tavares

Faculdade de Engenharia da Universidade do Porto, Portugal

maria.vasconcelos@fe.up.pt, smr@estsp.ipp.pt, dfreitas@fe.up.pt, tavares@fe.up.pt