MRI analysis for articulatory studies in Engineering and Philology Degrees using DICOMPAS Tool

J. L. García Arroyo[1], B. García Zapirain[2], Alex Iribar[3], Amaia Mendez[4], Ibon Oleagordia[5], R. M. Pagola[6], Itziar Turroz[7]

Abstract *— This work has been developed for engineering studies related with signal processing algorithms and for philology studies about the articulatory model of a language. By this work a study for the modelisation of the correct articulation of phonemes is presented, made through the treatment of MRI 2D images in the midsagittal plane. A software system was developed; implementing mathematical algorithms that allow the calculation of objective measures of articulatory model parameters obtained through treatment of images, and was used by researchers in linguists, phonetics and engineering. Engineering students can use image processing, statistic and software engineer techniques and even provide to the tool new screening algorithms for calculation of parameters, and the students of linguistics can make the scientific analysis of the characteristics of the language they are studying. The authors (in alphabetical order) are a multidisciplinary team consisting of researchers and professors.*

Index Terms — Image processing, Cine MRI, Articulatory parameters, Speech production, Vocal tract.

INTRODUCTION

University studies currently incorporate new information technologies, as a pedagogical tool in order to increase the competences the students should acquire in the degree they are pursuing. This article presents a software solution called DICOMPAS for practices of philology and engineering students, which allow them, using the different functionalities the tool presents, to deepen in the theoretical knowledge of the master classes in a practical way.

This software system is built for the study of the morphology of the vocal tract in the production of sounds and its changes in real time, over the analysis of MRI imaging sequences, consisting on the capture during a time lag of a sequence of images of magnetic resonances taken to informants while they are articulating the sound. There currently are images regarding sounds in Basque and in Spanish, though they can be used for other languages.

In such images, the software tool allows visualizing how the sounds are articulated, providing digital image processing utilities that allow observing the details and enables the measurements with the aim of characterizing them. Such measurements can be subsequently studied using statistical techniques.

DICOMPAS has been developed by a multidisciplinary working team of the University of Deusto, comprised by two working teams, one composed by phonetic researchers and teachers, who know this field of study and are the ones who guide the specification of the application, and the other one composed by engineering researchers and teachers, specialized in software engineering, mathematics, artificial intelligence and digital image processing, who designed and developed the algorithms which shape the software system.

Initially, it was built for the accomplishment of articulatory characterization studies, and in fact it has been used for several time in research works by phoneticians and engineers, having obtained important results in the study of Basque and Spanish sounds, leading in both languages, taking as examples the references in [1] y [2], which are in the line of other works like [3] y [4] for the study of English and Italian sounds respectively. The utility of this tool has been subsequently seen under the pedagogical perspective, and it is started to be used for the practices of philology and engineering students, with very satisfactory results.

On the one hand philology students are able, from the different sequences of images, to study the articulation of Basque and Spanish sounds, as well as its time-space characterization, over real speaker individuals. It is a pedagogical innovating tool, since, at least as far as the knowledge of the team is concerned, there are no similar tools in this discipline. The most widely used pedagogical device is the University of Iowa web page, which shows the articulation of sounds in various languages, through some pictures presented in a lively way [5]. It is a good resource for the study, nevertheless, it does not deal with real images, and furthermore, the student is not able to carry out any sort of interaction or measurement with the images.

On the other hand, the students of computer and telecommunications engineering are able to deepen in the

[1] J. L. García Arroyo, University of Deusto. Faculty of Engineering. Bilbao. Spain. jlgarcia@deusto.es
[2] B. García Zapirain, University of Deusto. Faculty of Engineering. Bilbao. Spain. mbgarciazapi@deusto.es
[3] Alex Iribar, University of Deusto. Faculty of Philology. Bilbao. Spain. airibar@deusto.es
[4] Amaia Mendez, University of Deusto. Faculty of Engineering. Bilbao. Spain. amaia.mendez@deusto.es
[5] Ibon Oleagordia, University of Deusto. Faculty of Engineering. Bilbao. Spain.ibruiz@deusto.es
[6] R. M. Pagola, University of Deusto. Faculty of Philology. Bilbao. Spain. rpagola@deusto.es
[7] Itziar Turroz, University of Deusto. Faculty of Philology. Bilbao. Spain. iturrez@deusto.es

DOI 10.14684/ICECE.8.2013.58-63

study of digital medical image processing techniques, being able to use in addition to practice statistical techniques through the measurements taken in them. Besides, they will be able to familiarize with software engineering leading techniques, or even include their own algorithms for the calculation of new parameters.

Additionally, it is considered important to note that this system allows the work among students from different degrees, something very enriching in this times of necessity of collaboration among professionals of different fields.

METHODS

MRI

MRI ("Magnetic Resonance Imaging") is a non invasive technique which uses the phenomenon of the magnetic resonance to obtain information about the structure and composition of the body to be analysed. It is frequently used in medicine, since it gives good results and it does not present radiation problems, as is the case of other techniques such as the X-rays. The information is captured by an MRI scanner and transformed into images of the inside of what has been analysed, in DICOM format, thereby generating the three-dimensional structure on the basis of those images, whose study is conducted using CAD ("Computer Aided Diagnosis") systems [6].

DICOM

DICOM ("Digital Imaging and Communication in Medicine") is the worldwide standard for the storage and transmission of medical images. DICOM files consist of a header and a body. In the header metadata of different kinds are stored of, among which the data of the patient, the date in which the study was conducted, the device used for the capture and the spatial resolution of the containing images are taken. In the body the proper images taken are stored, admitting different standards of compression format [7].

Cine-MRI

For the analysis of the articulatory model the most used technique, which is used in the present work, is the cine-MRI, consisting on the capture of a sequence of MRI DICOM images in a time lag. Such images can be either of 2D type (examples [8], [9] and [10]) and of 3D type (examples [11] and [12]). In this system 2D images of the midsagittal plane, preferred by the phoneticians for the obtaining of articulatory parameters [12], have been chosen, as can be seen in fig. 1 and 2.

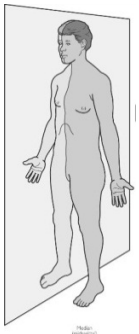

Figure 1: Midsagittal plane. Source in [13]

Figure 2: Example of cine MRI image

IMAGE DATABASE

6 native speakers of Spanish were selected for the tests in Spanish, and 6 native speakers of Basque for the tests in Basque; all of the individuals were adults, students and teachers of the University of Deusto, which previously training sessions were made with, in the phonetics laboratory.

Each one of the informants were subjected to a battery of questions, each one of them related to the articulation of a different sound, being 34 the total amount of sounds. In each one of the tests, both for the vowels and for the consonants, the individual articulates a sound during a time lag, capturing different instants of this action in a sequence of 24 phonemes of cine-MRI. The test design was made in such a way that, to the extent possible, it could be possible to capture the central moment of the corresponding sound the maximum number of times, facing to obtain the maximum number of samples concerning the articulation of such sound. To do so, attempts had been made to undertake the pronunciation of each sound in a normal phonetic position, and in a sustainable way along as possible.

The recording sessions of the MRI images were carried out in the Quiron Hospital of Biscay, GE 1,5T. –HealthCare– team obtaining dynamic sequences SSFSE (Single Shot Fast Spin Eco) with TR of 4, 8 and of 1, 3, 24 frames per second in a midsagittal plane.

SYSTEM DESIGN

In this section the system design is to be explained. To do so, in the first place, a high level view of the system will be given, secondly, the software tool will be presented, in the third place the main functionalities will be explained, and finally, in the fourth place, the main tasks the students are able to accomplish with this software solution will be summarized.

High-level view

A high level view of the system will be given thereupon:

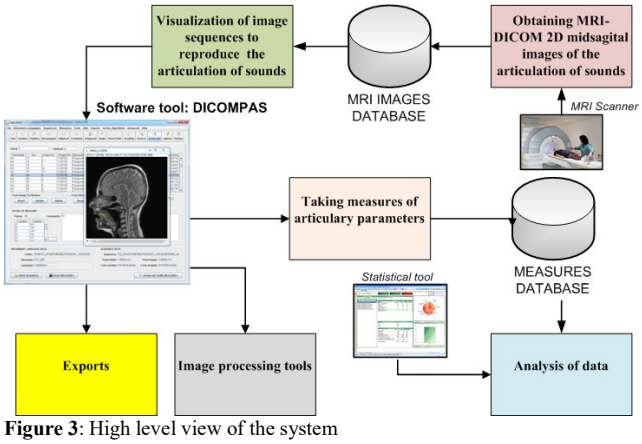

Figure 3: High level view of the system

As can be seen in fig. 3, the system is supplied by MRI images, taken from the MRI scanner, a task explained in the previous section, in which the MRI image database is described. Through these images, using the DICOMPAS software tool, the sound production process can be reproduced to a morphological level. It provides also with digital image operations for its analysis. It is also given the opportunity to take different measures in such images, facing two objectives. On the one hand to draw and identify visually certain interesting aspects of the articulatory process, and on the other hand to undertake an articulatory characterization of quantitative type, which is stored in a database of measurements. These measurements can be subsequently analysed using statistical techniques, using external tools like Excel y SPSS. Additionally, everything done can be exported to a wide range of formats, facing to be managed external tools.

Software tool: DICOMPAS

The DICOMPAS software tool is written in the Java programming language [14], using for the digital image processing the ImageJ framework [15]. For the reader to get a better idea of the appearance, in fig.4, the main screen of the application is shown. On such screen, a sequence of images of an informant in Spanish, pronouncing the "l" sound remains open. Over it several measurements are being taken, being introduced at this precise moment one of polygonal type, corresponding to the lingual reverse side.

The main functionalities of the tool, which have been commented before, are to be explained in the subsequent section. Beside these functionalities, which are the more relevant ones, the tool also provides basic management functionalities, as well as a great number of utilities to facilitate the work. Notwithstanding all of them shall not be explained in detail due to not deem interesting here.

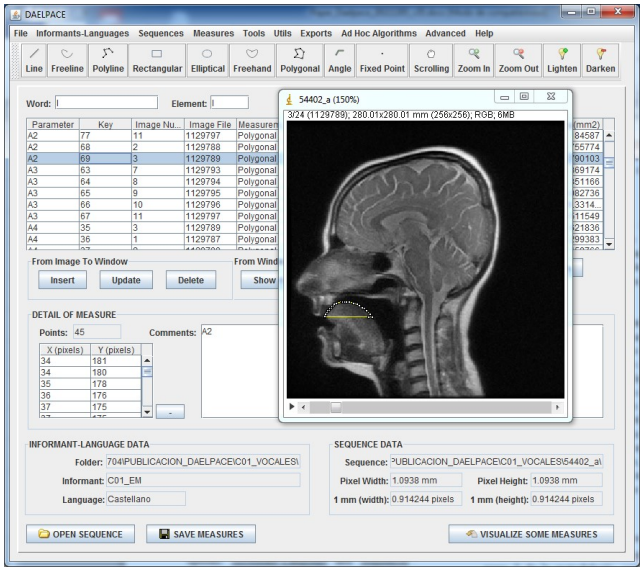

Figure 4: Main window of the software tool

Main Functionalities

Visualization of the MRI sequence of images

As it can be seen in fig. 4, using the DICOMPAS tool, the student is able to choose an informant of any language (Basque or Spanish), and a MRI sequence of images, corresponding to one sound, being able to visualize in detail the different parts of the anatomy of the vocal tract participating in the articulation of the sound, by means of the images. Additionally, its behaviour in the dynamic level can be observed, thanks to the possibility to generate the animation through such sequence.

Image Processing Tools

To facilitate the task of analysing the MRI images, the DICOMPAS software tool provides utilities for the student to be able to carry out digital image processing tasks such as the Zoom, Brightness, Edge Detection, Invert, Smooth, different types of linear and non-linear filters, and so on. Moreover, the engineering student is able to add digital image processing algorithms of their own production, carrying out image transformation, segmentation, object recognition operations, and so on.

Measurements

The student is able to take measures of different types over the images for the articulatory characterization, which are stored in the database. Such measurements can be accomplished with two objectives, on the one hand to draw certain regions of interest of the vocal tract while the sound is articulated, face to its visual identification, and on the other hand the obtaining of quantitative data concerning the articulatory parameters.

The different measurements that can be taken are:

Line: lines between certain relevant points (1 in fig. 5).

Freeline: curves taken freehand. Once they are saved, a

conversion process from freeline to polyline is carried out, being able to choose the number of points describing it. Thus, images saved of freeline type do not exist as such.

Polyline: curves determined by N points (2 in fig. 5).

Rectangular: regions with rectangle shape (3 in fig. 5).

Elliptical: regions with ellipse shape (4 in fig. 5).

Freehand: regions taken freehand. Once they are saved in a conversion process from freehand to polygonal is carried out, being able to choose the number of points describing it, the same way as the freehand case. Thus, measures taken of freehand type do not exist as such.

Polygonal: regions corresponding to polygons of N vertices (5 in fig. 5).

Angular: angles taken (6 in fig. 5).

Fixed point: They correspond to interest points (7 in fig. 5).

In fig. 5 can measures can be seen graphically:

Figure 5: 1 2 3 4 5 6 7

With respect to the properties captured for such measurements, for the ones of line and free line, the length is obtained, for the rectangular, elliptical and polygonal, the perimeter and area are obtained and for the angle the value of the angle is obtained. For the measurement of fixed-point type no property is obtained, it is used only to remark relevant points. In fig. 6 the measurements are illustrated with two examples:

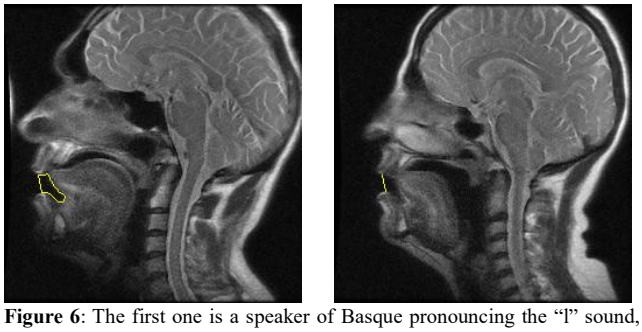

Figure 6: The first one is a speaker of Basque pronouncing the "l" sound, from which a measure of Polygonal type is taken. The second one is a speaker of Spanish, also pronouncing the "l" sound, from which a measurement of Line type is taken.

Analysis of Data

From the measures taken the student must carry out the analysis using statistical techniques and tools like Excel or SPSS. In this way, the student might be able to draw some conclusions from the quantitative point of view, something especially important for those who wish to dedicate in the future to research works.

Exports

Everything accomplished can be exported to a wide range of formats, face to be managed by external tools.

The DICOMPAS software tool allows exporting the different values of the measurements to different formats, with different selection and data arrangement criteria.

XML: export of the measurements to XML for the interchange of information with other platforms. XML-Schema syntax is followed, which is defined in this Project.

RTF: export of the measurements to de RTF format.

Excel: export of the measurements to Excel, which enables the accomplishment of statistical analyses, both with Excel and with other data analysis tools like SPSS.

JPG: In JPG format, the measurements, either alone, or drawn over the images, either one or more than one, with the aim of being able to be consulted graphically.

Summary of the main tasks that may be performed by students

Hereby, with the purpose of studying the procedural learning objectives obtained from the pedagogical point of view, a summary with the main tasks allowed to be performed by philology and engineering students using this software solution is carried out.

PHILOLOGY STUDENTS

- The study of the different parts of the anatomy of the vocal tract that intervene in the speech production.
- Visualization of the production of the sound through the animation of the sequence of images.
- The possibility of taking measures concerning certain parts or aspects of interest for the study of the articulatory characterization.
- To be able to carry out scientific works through the measures taken, using statistical techniques.

ENGINEERING STUDENTS

- To be able to manipulate real MRI medical images, in the DICOM standard, being able to prove different digital image processing techniques.
- Possibility of undertaking statistical studies concerning the measures taken.
- Being able to develop algorithms were they could implement their knowledge of software engineering (Java, database, XML, etc.) and digital image processing.

RESULTS

The proposed system has been used with two groups of Basque and Spanish philology students and 4 groups of telecommunications and computing engineering, with excellent results. In fig. 7 a student using the DICOMPAS software tool is shown:

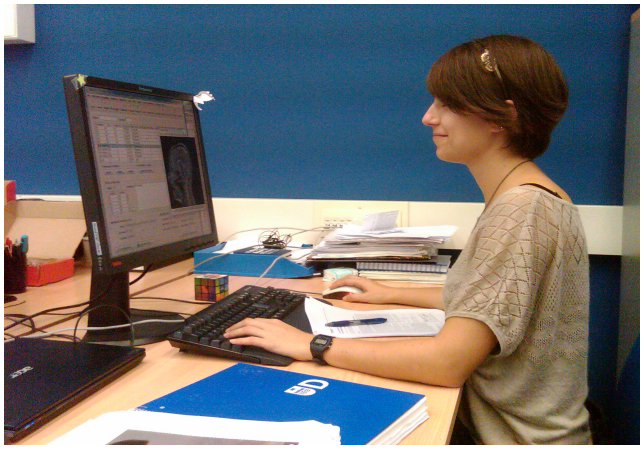

Figure 7: Student using the DICOMPAS tool.

In the case of philology students the experience was satisfactory, since the theoretical concepts they had the opportunity to study about the vocal tract were able to be applied to examples of images corresponding to real speakers, thus checking in a practical manner what has been explained so far in the theory. Furthermore, detailed measures were able to be taken of different interesting aspects, thus being able to deepen in the articulatory description and, in addition, they were able to conduct brief research projects concerning articulatory characterization, obtaining quantitative data and undertaking over such information statistical inferences, which trained them in methods usually used in scientific researches.

In the case of engineering students it was also successful; they were able to practice digital image processing techniques, in the standard format used in medicine. Additionally, statistical techniques were able to be implemented through the data obtained from the measures taken; applying them in real cases, in this particular case aimed at the accomplishment an articulatory characterization of quantitative type. They were also able to familiarize with the techniques of software engineering used in this tool, such as the java language, databases, XML files, and, of course, about the digital image processing over the ImageJ API, having the possibility for the developing of their own algorithms added over the platform.

The students of both fields were also able to collaborate in a interdisciplinary working experience, contributing each one with their knowledge, which is considered of great importance since it is currently of great necessity the collaboration among experts in different fields.

Finally, it was observed that this experience produced a strong motivation among the students towards the research work, as they could undertake in a practical manner several works similar to those real researches using this tool.

CONCLUSIONS AND FUTURE WORK

DICOMPAS was presented; an innovative software system for the accomplishment of phonetic studies concerning articulatory characterization of the Basque and Spanish sounds, over cine-MRI sequences of images of 2D, type, taken in the midsagittal plane, in native speakers of Spanish and Basque. This system has been developed by a multidisciplinary teamwork from the University of Deusto, composed by two research groups, coming from the fields of philology and engineering.

From the pedagogical point of view it was checked the usefulness of this system, in fact, it is starting to be used with philology and engineering students, with very satisfactory results measured hitherto.

The students of philology are able to increase their knowledge of the articulatory characterization in the production of a sound and they are also able to conduct measurements and statistical studies for the obtaining of quantitative data, whereas the students of engineering are able to improve to a practical level their knowledge digital image processing, software engineering and statistics.

The system, which is built in a modular manner, is very scalable, and in the future it will continue evolving. It will continue to be used as a pedagogical tool, studying the interaction of the students with the system, and improving it with a view to the enhancement of its learning. More MRI images will be added to the image database, new algorithms and functionalities will be developed, and new articulatory characterization studies will continue to be conducted.

ACKNOWLEDGMENTS

This research work is partly funded by the Basque Government Department of Education and Research, as well as, to the University of Deusto for its support.

REFERENCES

[1] R. M. P. Ana Elejabeitia, Alexander Iribar, “El cine-MRI aplicado a la descripción de las sibilantes vascas,” *Estudios de fonética experimental*, vol. 18, 2009.

[2] I. T. Alexander Iribar, R. M. Pagola, “Observaciones sobre la articulación de la lateral alveolar en euskara y castellano,” in *V Congreso de Fonética Experimental, Cáceres*, 2011.

[3] X. Zhou, “An MRI-based articulatory and acoustic study of American English liquid sounds/r/and/l,” *PhD dissertation. University of Maryland, College Park*, 2009.

[4] A. Romano and P. Badin, “An MRI study of the articulatory properties of italian consonants,” *Estudios de fonética experimental*, vol. 18, 2009.

[5] University of Iowa, “Phonetiks.” [Online]. Available: http://www.uiowa.edu/~acadtech/phonetics/.

[6] J. Hornak, *The Basics of MRI*. ScientificCommons, 2002.

[7] NEMA, “DICOM specification.” [Online]. Available: http://medical.nema.org/dicom/.

[8] P. Ladefoged, “Generating vocal tract shapes from formant frequencies,” *The Journal of the Acoustical Society of America*, vol. 64, no. 4, p. 1027, Oct. 1978.

[9] D. Beautemps, P. Badin, and G. Bailly, “Evaluation of an articulatory-acoustic model based on a reference subject,” *ESCA Tutorial and Research Workshop on Speech Production Modeling - Speech Production Seminar*, 1996.

[10] D. Karthikeswaran and S. Dinakar, “Developing a scientific visualization tool for Inner articulators,” in *2011 International Conference on Emerging Trends in Electrical and Computer Technology*, 2011, pp. 480-488.

[11] P. Badin, G. Bailly, L. Revéret, M. Baciu, C. Segebarth, and C. Savariaux, “Three-dimensional linear articulatory modeling of tongue, lips and face, based on MRI and video images,” *Journal of Phonetics*, vol. 30, no. 3, pp. 533-553, Jul. 2002.

[12] P. Badin and A. Serrurier, “Three-dimensional modeling of speech organs: Articulatory data and models,” *Trans. Tech. Comm. Psychol. Acoust., The Acoustical Society of Japan.*, vol. 36, no. 5, H–2006-77, pp. 421-426, Jul. 2006.

[13] “The free dictionary.” [Online]. Available: http://medical-dictionary.thefreedictionary.com/midsagittal+plane.

[14] Sun Microsystems, “Java.” [Online]. Available: java.sun.com.

[15] U.S. National Institutes of Health, “ImageJ.” [Online]. Available: http://rsbweb.nih.gov/ij/.