FCT project: PTDC/SAU-BEB/104995/2008

Title: Assistive Real-Time Technology in Singing

Minutes of the First Plenary Meeting of the Project

November 19th 2010

This metting took place in room I-105 at the Faculdade de Engenharia da Universidade do Porto (School of Engineering of the University of Porto) and has gathered the following project members and institutions:

Aníbal Ferreira (FEUP)
Pais Clemente (FMUP)
Álvaro Barbosa (UCP)
Sofia Serra (UCP)
Ricardo Sousa (FEUP)
Sandra Dias (FEUP)
Rui Taveira (ESMAE-IPP)
Sten Ternström (KTH)

Professor Sten Ternström has kindly accepted the invitation to give a seminar on the day before the meeting. The title of the seminar was "Does the acoustic waveform mirror the singing voice ?". In addition to several members of project, the seminar has attracted more than 15 people not familiar with the project and not known to the project coordinator. Some of them have contacted the project coordinator with the purpose to receive more information about the project. The seminar consisted of a presentation regarding the KTH research group in the area of musical acoustics and its research activities, and a very inspiring presentation on a perspective of the different layers involved in the communication chain of singing. The presentation was clearly acknowledged by most of the attendees.

The purpose of the plenary meeting on Nov 19th was to discuss the main tasks the project and their realization objectives, the project members involved, budget and immediate actions. Two PDFs addressing these aspects have been released to the project members previously to the meeting.

Regarding Task 1
Name: Correspondence between subjective quality parameters of the singing voice and objective acoustic features

This task has generated a rich discussion on the scope and ambition of underlying objectives. There was a general opinion that the expected results for this task were too

ambitious for the available time frame. It was suggested that a limited number of perceptual quality parameters in singing be first selected to establish a correspondence with objective features amenable to a useful visual representation on screen. A discussion of the nature of those parameters in terms of "quality" or "correctness" is to be clarified based on actual practice. An agreement has also been established that the project will focus essentially on lyrical voice. Suggestions were also made to make available to this task the studies and results already obtained by Susana Freitas in the context of her PhD dissertation involving the correspondence between acoustic features and perceptual quality parameters of normal/pathological voice. An MSc. student from FEUP has already been motivated to this task. He should work with an MSc. student recruited from ESMAE-IPP. The announcement of this scholarship must be immediate.

Other facts, remarks and progress so far
The PhD dissertation of Susana Freitas is strongly related to this task as it involves the correspondence between acoustic (objective) features of normal/pathological voices and perceptual (subjective) quality parameters. Right now she is developing her PhD work in Brasil but she will be back to Porto in January 2011. The opening of BM1 and BD1 scholarships must be immediate.

Regarding Task 2
Name: New technology-assisted methodologies in singing teaching/learning

This task involves one of the main realizations of the whole project which is the interactive software environment assisting teachers and students in singing teaching, learning and training. It was stressed during the discussion that the design of all the functionalities of the software should be driven by the practitioner. In this perspective and as planned in the project description, a first functional version of the software should be made available to the ESMAE-IPP and UCP partners as soon as possible so as to evolve by incorporating feedback by users. Prof. Álvaro Barbosa has indicated that the UCP is willing to make a proposal of a suitable hardware selection (involving the acoustic signal acquisition, the touch screen display and the sound reproduction equipment) to be used at all three sites (ESMAE-IPP, FEUP and UCP). A discussion was also made on the impact of the relation between the teacher and the student. It was clarified that the main intention of this task is to provide a tool which tries to simplify, to facilitate and to make more effective the communication between the student and the teacher, not to replace the teacher in any way. Profª Sofia Serra has informed that her PhD dissertation, which she completed in June this year, was about the teacher-student interaction and relationship. In this perspective, this task also motivates the topic of computer-mediated human-to-human interaction which is in line with the main concerns of the project and task (hence the strong involvement of Prof. Rui Taveira). A remark was also made that the project description appears to determine already the main functionalities of the software environment. It was clarified that the task description addresses a number of potential functional features but the ultimate selection and specification will be dictated by real users.

Relevant facts, remarks and progress so far
Prof. Rui Taveira has started already his PhD program at FEUP on Digital Media.

Regarding Task 3
Name: Singing to musical score transcription and music composition

This task has generated a discussion regarding the balance between research and development since the latter appears to dominate and the available base-line software already includes the basic resources for singing-to-MIDI transcription. It was recognized that this task is essentially development oriented (by including simple editing capabilities such as note removal, insertion, normalization and modification) but also includes opportunities for research. For example, the process of singing to MIDI interpretation should be "informed" and flexible enough such as to avoid that vibrato is converted into a fast discrete note oscillation. Research opportunities also exist in the perspective of the use of the software platform as an interactive mouse-free and keyboard-free tool for educational composition modalities that are fitted to workshops at Casa da Música, as proposed in the task description. Other possible uses of the music-score functionality is as an interactive, real-time, score-follower tool as it exists in commercial products such as "StarPlay" and "Music Master Works". This is to be decided later on.

Relevant facts, remarks and progress so far
An MSc student from FEUP has chosen the topic of this task for his dissertation. Work will start immediately with the study of the state-of-the-art and with code familiarization with the support of BD1.

Regarding Task 4
Name: Correspondence between objective acoustic features of the singing voice and voice disorders in singing

This task is similar to Task 1 in the sense that it focuses on the correspondence between objective acoustic features and possible disorders in singing. It has been pointed out by Prof. Sten Ternström that singers tend to be good at camouflaging singing voice problems so databases of singing voices with problems or disorders are likely to be hard to gather. Prof. Ternström indicated that a similar topic has already been addressed at KTH in the context of a PhD dissertation by Anick Lamarche (http://www.speech.kth.se/~anick/). This dissertation has produced an annotated database of singing voice where the singer has signaled (by means of a manual marker or button) the regions in singing where discomfort or reduced vocal control was felt. An analysis of the audio files, with the purpose to identify features correlating with the discomfort, has however not yet been done. The phonetogram was used instead to identify regions of interest of the singing voice showing problems. It has also been pointed out that it could be more helpful to look for "late correlates" of the problem such as loss of high notes. Since the work performed by Anick Lamarche is very important to this task, Prof. Sten Ternström will assess if it is possible to share the existing database. Other database possibilities are indicated in the

task description. Extensive contacts with singing professionals in Portugal are likely to be established with the purpose to collect singing voice material including evidence of voice problems.

Relevant facts, remarks and progress so far
The opening of the BD3 scholarship (by FMUP) must be immediate.

Regarding Task 5
Name: Robust real-time glottal pulse estimation from running singing

This task has generated some discussion regarding the pertinence of inverse filtering, the validity of the source-filter model assumption and, especially, the difficulties that are likely to be faced due to the feedback interaction between the source and the filter. The project coordinator has indicated that experiments were planned, with the help of an ORL doctor, to collect acoustic data reflecting the glottal flow (using a tiny microphone attached to a video-laringoscope probe) and simultaneously recording the voice signal outside the mouth. This acoustic data will provide some opportunities to study the effect of the vocal tract filter and, desirably, to help the design of a new frequency-domain approach to inverse filtering, as described in the task description.

Relevant facts, remarks and progress so far
Research work has already been developed and published regarding this task, concerning in particular the estimation of the relative delay of partials pertaining to the harmonic structure of a spoken or sung vowel. It is believed that this result may pave the way for a new frequency-domain approach to glottal pulse estimation in real-time and from running singing. A second-cycle MSc student (Sandra Dias) has selected the topic of this task as her dissertation topic. This work is to be articulated with the expertise of BD3 to be recruited by FMUP.

Regarding Task 6
Name: Real-time preventive assessment of the singing voice

This task is meant to take advantage of the results obtained in the context of tasks 4 and 5. It has been acknowledged that nowadays there is a clear receptivity to the idea of preventive voice risk assessment which reinforces the pertinence of this task. In this perspective, it makes sense to first look at what has and can be done with speech (before singing is addressed). Prof. Sten Ternström has noted that acoustic facts should be taken into consideration such as the case of the spectral slope saturation when the loudness increases. Also, as already indicated above concerning task 4, when there is a problem with the singing voice, it first happens that the singer feels discomfort but that may not be evident in the sound and thus may not be perceived by the listeners. This should be taken into consideration in this task as well as the fact that singers may deliberately be singing at altered (and therefore more stressful) tones in order to create specific impressions.

Relevant facts, remarks and progress so far
Because this task is the natural continuation of task 4, the opening of the BD3 scholarship (by FMUP) is urgent and must be immediate.

Regarding Task 7
Name: Management

Immediate announcements for scholarships BM1, BD2 and DB3 has been discussed. The project coordinator has also informed that the first money transfer (20% of the budget for each partner institution ESMAE-IPP, FMUP, and UCP) would take place within the next few weeks.

Finally, the project coordinator thanked all members for the presence and participation. Partners have acknowledged the fruitful discussions during the meeting and the meeting was adjourned.

AJF, November 19th, 2010