Post-doctoral position (6 months) – GIPSA-lab, Grenoble, France

Assistive speech technology based on ultrasound imaging

Duration, location and staff
The position is open from 2014 September (until filled) for a duration of 6 months. The work will take place in GIPSA-lab, Grenoble, France (in collaboration with TIMC-IMAG lab).

Researchers involved: Thomas Hueber, Pascal Perrier, Pierre Badin (GIPSA-lab), Yohan Payan (TIMC-IMAG).

Context
Within the last four decades, ultrasound imaging has been used in phonetics to visualize tongue movements during speech production. Ultrasound imaging is a non-invasive and clinical-safe technique providing vocal tract scans in the mid-sagittal or coronal plane, with good spatial and temporal resolution. More recently, ultrasound imaging has also been used in the context of assistive speech technologies. In [1], we showed that ultrasound can be used to drive a speech synthesizer. Such a system was called a “silent speech interface” since it allowed speech communication without the need to vocalize the sound. It can potentially supplement a production system no longer functional, for instance after a massive surgery such as a total laryngectomy. In recent years, the use of ultrasound imaging for speech therapy has also received considerable attention. Ultrasound can be used as a biofeedback tool that helps a patient to correct an articulation disorder by visualizing his/her own tongue gesture [2]. However, raw ultrasound images might be difficult to “read” for a naïve speaker. It is plagued by a typical noise called speckle and does not show the limits of the oral cavities, i.e. neither the palate nor the pharyngeal wall.

The « Living Book of Anatomy » (LBA) project, supported by the Persyval Labex, aims at developing a series of systems allowing a user to visualize the underlying biomechanical mechanisms of a human body in action, using augmented reality. These tools could have several applications, such as the learning of anatomy by medical students, or the design of immersive rehabilitation protocol of physical or cognitive disorders. The underlying assumption of this work is that experiencing the anatomical knowledge with his/her own body, or embodiment, could make the learning of anatomy and the rehabilitation easier.

Project description
In the context of speech therapy, we aim at designing a visual feedback tool showing a patient his/her tongue movements, captured by ultrasound imaging, but displayed in an intuitive manner. Two scenarios will be envisioned (as shown in the figure). In the first one, the ultrasound images will be “augmented” by overlaying the tongue contour tracked in real-time (as shown in Figure 1). In the second one, the extracted tongue movements will be used to drive a 3D biomechanical model (such as the one developed by Buchaillard et al [3]). This approach could make the visualization very intuitive, since it displays all the internal structures of the vocal tract (such as the palate and the pharyngeal wall). To that purpose, the goal of the recruited post-doctoral research fellow will be twofold:

Developing robust segmentation technique for automatically tracking the tongue in ultrasound images, in real-time. These techniques could be based on region-based (region growing), or model-based (ASM, AAM) techniques. Developed technique will be implemented (in C++) on the Ultraspeech platform (www.ultraspeech.com).

Designing robust pattern recognition techniques for extracting sequences of elementary tongue gestures in ultrasound image stream (such as phonetic or sub-phonetic gestures). The decoded gesture will be used in the “Living book of anatomy” to trigger animation of the 3D biomechanical model. These techniques could be based on Hidden-Markov Model (or more complex graphical models), Conditional Random Fields, or Time-delay Neural Network.

Keywords: ultrasound image segmentation, robust image processing, machine learning, speech therapy, augmented reality, biofeedback, embodiment, real-time.
Prerequisite: PhD degree either in image processing or in machine learning, or in audiovisual speech processing. A background in medical image processing (including recent segmentation techniques) or graphical models is a plus.

To apply: Applicants should email a curriculum vitae along with a brief letter outlining their research background, their list of references (at least 2), and a copy of their two most important publications, to Thomas Hueber (thomas.hueber@gipsa-lab.fr) and Pascal Perrier (pascal.perrier@gipsa-lab.fr).

Salary: The gross salary per month is ~2300€

References: