ECOLE DOCTORALE EEATS

Electronique, Electrotechnique, Automatique, Traitement du Signal

---

PROJET DE THESE POUR ALLOCATION FLECHEE PAR LE MINISTERE
Version en anglais

Doctoral school: EEATS .....

Doctoral school number: 220 .....

Laboratory: GIPSA-lab, Speech & Cognition Dept.

Label **: .....**

Address **: ENSIEG, Domaine Universitaire - BP46, 38402 Saint Martin d'Hères**

Thesis supervisor: Pierre Badin

e-mail address: *Pierre.Badin@gipsa-lab.inpg.fr*

Co- supervisor(s): Gérard Bailly

Title: Control of talking heads by multimodal inversion – Application to language learning and rehabilitation

Context and problem :

Speech production necessitates fairly precise control of the various orofacial articulators (jaw, lips, tongue, velum, cheeks, etc.). Regulating these gestures implies that a fairly precise feedback about his / her vocal production is available to the speaker. Auditory feedback is essential and its degradation can generate degradation, if not total loss, of speech production capabilities. In fact, the perception of the acoustic consequences of articulatory gestures can be degraded in different ways: either peripherically through the degradation, if not the complete loss, of this feedback (deaf and hearing impaired people, implanted or not), either in a more central way through the loss of sensitivity to phonological contrasts due to phonological deafness (contrasts not exploited in the mother language: i.e. Japanese speakers have extreme difficulties producing the /l/ vs. /r/ contrast not exploited in their mother language).

The stake of this doctoral work is to explore the speakers' abilities to exploit a virtual multisensory feedback that complements, if not substitutes for, the failing auditory feedback. The virtual feedback that will be designed and studied in this framework will be provided by a talking head (see on the right in 2D or 3D) that reproduces in an augmented reality mode – in real time or offline – the articulation of a sound for which only the acoustical and / or visual signal is available.

The thesis challenge is to design and assess a robust system that can estimate the articulation from its sensory consequences and in particular that deals with the normalisation problem (establishing the correspondence between the audiovisual spaces of the talking head and of the speaker), and then to quantify the benefit that an hearing impaired person or a second language learner can gain from a restored sensory motor feedback loop.

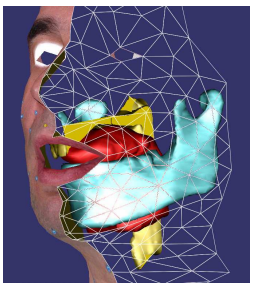
ECOLE DOCTORALE EEATS

Electronique, Electrotechnique, Automatique, Traitement du Signal

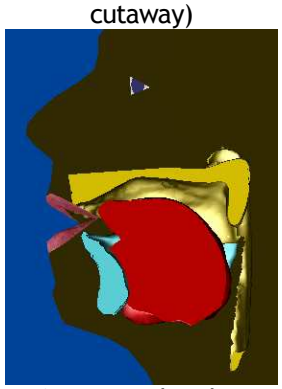
---

Work environment:

The MPACIF (Talking Heads, Communicating Agents, Face to Face Interaction) team of the Speech and Cognition department of GIPSA-lab has developed a full articulated virtual talking head from the analysis and modelling of the geometrical deformation of the visible and not visible organs induced by orofacial muscles. The technical work is mainly centred on the control of articulation for a target language using statistical learning methods (GMM, HMM, Bayesian networks, etc.) recently introduced in the domain of inversion, trajectory formation and of multimodal speech synthesis.



Augmented reality rendering of articulation (3D cutaway)



Augmented reality rendering of articulation (2D cutaway)

Stake :

This research work proposes three important challenges: basic research on multimodal perception and perceptive substitution, more technological research on the design and assessment of technical solutions for the learning techniques of sensori-motor maps, and an application stake in the domain of speech rehabilitation or learning.

Contacts :

Pierre Badin GIPSA-Lab 04 76 57 48 26 Pierre.Badin@gipsa-lab.inpg.fr
Gérard Bailly GIPSA-Lab 04 76 57 47 11 Gerard.Bailly@gipsa-lab.inpg.fr

Knowledge and skills required:

The ideal candidate will have working knowledge in the following domains:

- statistical learning (HMM, Bayesian networks)/ automatic speech recognition
- speech perception and production in hearing impaired subjects
- man-machine interfaces and multi-modal communication
- audiovisual talking heads