Ultrasound study for a non-vocalized speech interface for medical and telecommunication applications

Mots clés:
- Directeur de thèse : Bruce Denby
- Co-encadrant(s):
- Unité de recherche : Laboratoire SIGnaux, Modèles, Apprentissage statistique
- Ecole doctorale : École Doctorale Informatique, Télécommunications, Électronique de Paris
- Domaine scientifique principal: Divers

Résumé du projet de recherche (Langue 1)

In the past few years there has been significant interest in a new type of speech recognition system using sensors that can give direct information on the movement of the articulators (essentially the tongue and lips), for example, ultrasound imaging or electromyographic sensors. Such systems can restore the ability to speak for persons who have lost the use of their vocal chords due to cancer or injury, and can also enable speech recognition in very noisy environments where the acoustic signal is corrupted. A third very interesting possibility is the so-called Silent Speech Interface, which would allow a user to engage in spoken communication without activating his or her vocal chords, thus providing a very secure and unobtrusive means of verbal communication in a variety of everyday situations; cellphone manufacturers are in particular very interested in this last possibility. At the Sigma Lab, ESPCI ParisTech, we have for a few years now been developing a Silent Speech Interface based on real-time ultrasound imaging of the tongue and video imaging of the lips during speech, and we have recently begun to obtain some very interesting results. The purpose of the the PhD thesis will be to go back to the physical fundamentals of ultrasound imaging (as well as perhaps including other types of sensors) for the vocal tract, and, rather than using ad-hoc medical diagnosis techniques, devise a new system specifically adapted to speech applications, in order to extract the maximum possible amount of acoustically relevant information from the ultrasound and other sensor signals. The goal will then be to move towards a high-performance and genuinely practical system, having applications in a wide range of areas both in medicine and in industry.

Résumé du projet de recherche (Langue 2)

The thesis will involve work in the following fields, among others: - ultrasound physics - medical imaging - image processing - signal, audio, and speech processing - machine learning (to learn the mapping between sensor signals and acoustic information) - electronics - prosthetic design This interdisciplinary mix will allow the student to gain competence in a wide range of fields valuable for a future career in signal processing and real time applications, and includes inroads to medical instrumentation and telecommunications research as well.

Informations complémentaires (Langue 1)

Devices capable of performing real time non-vocalized speech recognition for medical and telecommunications applications, using a variety of technologies, are currently being studied in projects in France, Germany, England, Japan, and the United States. Sigma lab, as a pioneer in the domain, enjoys a privileged position at the crossroads of this exciting, emerging new interdisciplinary field of research.

Informations complémentaires (Langue 2)