

# AVIS DE SOUTENANCE DE THÈSE DE DOCTORAT

**Monsieur OLIVEIRA MESSINA Ronaldo** soutiendra une thèse  
**Le Lundi 5 Décembre 2005 à 11H30**

---

**Amphithéâtre I.U.P  
AGROPARC**

SPÉCIALITÉ : **Informatique**

Titre de la thèse : *Modélisation acoustique pour la reconnaissance de la parole :  
unités longues et multi modélisation*

Membres du jury :

**Mr. DE MORI Renato**, professeur, Informatique, Université d'Avignon et des Pays de Vaucluse, Avignon

**Mr. DUPONT Stéphane**, Directeur de Recherche, Laboratoire Speech Processing Department, Traitement du signal, MULTITEL

**Mr. JOUVET Denis**, Directeur de Recherche, Laboratoire : SSTP/RVA, Traitement du signal, France Télécom

**Mr. LAFACE Pietro**, Professeur, Laboratoire : Dip Di Automatica e Informatica, Informatique, Politecnico di Torino

**Mr. WELLEKENS Christian**, Professeur, Laboratoire : Multimédia, Traitement du signal, Institut Eurécom

Résumé de la thèse :

Le domaine de cette thèse est la modélisation acoustique pour la reconnaissance vocale. Deux aspects principaux sont abordés : la modélisation par unités longues (groupes consonantiques, quasi-syllabes, multi-phonèmes) et la multi modélisation (c a d. comment « construire » efficacement une modélisation par classe de données) ?

Nous avons étudié aussi des méthodes pour accélérer l'apprentissage des modèles avec des alignements fixés, et aussi comment fabriquer des mélanges de Gaussiennes sans passer par les étapes de division usuellement utilisées pour construire les mélanges.

Il y a un intérêt croissant pour la modélisation d'unités plus longues que les phonèmes parce qu'elles offrent la possibilité de modéliser des dépendances temporelles de longue durée et des variantes de prononciation au sein du groupe de phonèmes considéré. Les unités longues, au travers de la modélisation contextuelle des groupes de phonèmes, ont donné de gains légers par rapport aux phonèmes. Lorsque les modèles sont adoptés à une application donnée, les unités longues sont significativement meilleures que les phonèmes.

Avec la multi modélisation, l'idée est d'avoir des modèles spécialisés pour différentes conditions, et donc plus précis pour chaque condition considérée. Un modèle spécifique à chaque condition est estimé à partir des données sélectionnées en fonction des connaissances a priori, par exemple un groupe de locuteurs (sexe, âge, accent), un type de canal (RTC, GSM), une tranche de valeurs de rapport signal à bruit, ou autre aspect du signal. Différentes méthodes pour combiner les modèles spécifiques à chaque classe sont comparées, et ce pour différents choix de connaissances a priori. Avec des modèles dépendants du sexe du locuteur et du canal de communication, il a été possible d'améliorer les performances par rapport à des modèles appris sur les mêmes données mais sans prendre en compte les classes de variabilité. Une combinaison au niveau acoustique (mélange de Gaussiennes) a donné les meilleures performances.