

**AVIS DE SOUTENANCE DE THÈSE DE
DOCTORAT**

Monsieur Sylvain MEIGNIER soutiendra une thèse
le Jeudi 14 Novembre 2002 à 14 heures

—
Amphi ADA LOVELACE

IUP GMI - Université d'Avignon et des Pays de Vaucluse

SPÉCIALITÉ : Informatique

Titre de la thèse : Indexation en locuteurs de documents sonores : Segmentation d'un document et Appariement d'une collection.

Membres du jury :

M. Jean-François BONASTRE, Maître de Conférence, HdR, Laboratoire d'Informatique (EA 931 FRE 2487), Université d'Avignon et des Pays de Vaucluse.

M. Henri MÉLONI, Professeur, Laboratoire d'Informatique (EA 931 FRE 2487), Université d'Avignon et des Pays de Vaucluse.

Mme Régine ANDRÉ-OBRECHT, Professeur, IRIT, Université Paul Sabatier, Toulouse.

M. Éric MOULINES, Professeur, Département TSI/CNRS URA 820, École Nationale Supérieure des Télécommunications (ENST), Paris.

Mme Catherine BERRUT, Professeur, Laboratoire CLIPS – IMAG, Grenoble.

M. Frédéric BIMBOT, Chargé de Recherche, IRISA/INRIA, Rennes.

M. Jean-Claude JUNQUA, HdR, Panasonic Speech Technology Lab (PSTL), Santa Barbara, USA.

Résumé de la thèse :

L'archivage sous forme numérique des documents offre beaucoup d'avantages. Il permet, en particulier, le transfert rapide et peu coûteux de l'information et le stockage de grandes quantités de documents dans des espaces physiques restreints. Cette accumulation de documents de natures variées crée un besoin d'accès efficace et rapide à l'information.

Cette thèse s'inscrit dans le cadre de l'indexation automatique de documents sonores en locuteurs. L'indexation en locuteurs permet de détecter les locuteurs présents dans un document ou dans une collection de documents et d'annoter le contenu des documents suivant ces locuteurs. Elle met à la disposition des moteurs de recherche un critère supplémentaire, leur permettant d'accéder aux annotations issues de l'interprétation des documents.

Deux thèmes ont retenu notre attention : la segmentation en locuteurs d'un document sonore et l'appariement en locuteurs d'une collection de documents sonores. La segmentation en locuteurs de documents détermine le nombre de locuteurs d'un document et les interventions de chaque locuteur.

L'appariement en locuteurs d'une collection de documents détecte le nombre de locuteurs présents dans la collection et classe les interventions par locuteur. Ce travail est réalisé à partir de documents préalablement segmentés (tâche de segmentation vue précédemment).

Dans le cadre de cette étude, nous proposons une approche au problème de la segmentation en locuteurs par une méthode itérative originale. Celle-ci repose sur une modélisation de la conversation par un modèle de Markov caché qui évolue à chaque détection d'un nouveau locuteur. La méthode proposée a obtenu des performances à l'état de l'art lors des évaluations internationales NIST en segmentation en locuteurs, tout en réduisant d'environ deux tiers les temps de calcul par rapport aux méthodes concurrentes.

L'appariement en locuteurs est une tâche récente et encore peu abordée dans la littérature. Dans ce travail, nous avons choisi une méthode fondée sur la classification hiérarchique ascendante. Nous proposons deux nouvelles mesures permettant de tirer profit des informations apportées par la segmentation des documents, ainsi qu'une nouvelle méthode de détection du nombre de locuteurs. Cette dernière repose sur l'élagage de l'arbre de classification sans nécessiter une coupe préalable de l'arbre. L'appariement de locuteurs a été évalué sur un corpus issu des évaluations NIST.