

Reconnaissance des Formes et Ingénierie des Technologies Vocales

Responsable : R. André-Obrecht (obrecht@irit.fr, 05.61.55.68.86)

Objectifs

Il est reconnu que le moyen le plus naturel de communication entre hommes est la parole et que la parole doit être considérée comme un moyen de communication homme-machine privilégié. Les systèmes automatiques de reconnaissance ou de synthèse de parole ont acquis des performances telles que leur intégration dans des systèmes interactifs est devenue effective. Le cours d'ingénierie des techniques vocales a pour but de donner les fondements théoriques du traitement automatique de la parole. La problématique des systèmes interactifs est abordée au travers de la compréhension automatique de la parole et des processus de gestion de dialogue.

Le cours de reconnaissance des formes a pour but de donner les connaissances de base pour traiter un problème relevant de ce thème. Les domaines d'application sont le traitement d'image (reconnaissance de visages, de caractères manuscrits, etc.) et le traitement de la musique et de la parole (reconnaissance de notes, d'instruments, du locuteur, ou de la langue).

Contenu (C : 46 – TP : 30)

1. Ingénierie des technologies vocales

- 1.1. **Initiation à la parole** : Production et perception de la parole, description acoustique des sons
- 1.2. **Bases en traitement du signal de parole** : Transformées de Fourier et analyse spectrale pour la reconnaissance, transformées pour le codage (LPC et type MP3)
- 1.3. **Modélisation statistique pour le traitement automatique de la parole** : Mélange de lois gaussiennes, modèles de Markov Cachés, modèles de langage (ngram, nclass)
- 1.4. **Les problèmes d'apprentissage** : Algorithme de type EM, adaptation de type MAP
- 1.5. **Méthodes d'évaluation** : Corpus, Ressources, Mesure de confiance. Fiabilité. Robustesse.
- 1.6. **Synthèse vocale** : Synthèse à partir du texte, synthèse par unités élémentaires (diphones...)
- 1.7. **Serveurs vocaux interactifs** : Architectures et applications, conception et développement d'un système de dialogue oral, compréhension de la parole et gestion de dialogue, évaluation des systèmes de dialogue oral

2. Reconnaissances des formes

- 2.1. **Analyse des données**: Quantification vectorielle, analyse en composantes principales, analyse factorielle
- 2.2. **Classification Discriminante**: cas des classes linéairement séparables, méthodes non linéaires, support Vector Machines
- 2.3. **Réseaux de neurones** : perceptron multicouches, cartes de Kohonen
- 2.4. **Approche Bayésienne Bayesian approach** : méthodes paramétriques et non paramétriques
- 2.5. **Apprentissage et adaptation** : algorithme EM, algorithme du Maximum a posteriori

Pré-requis

Aucun.

Bibliographie

- *La parole et son traitement automatique*, Calliope, Masson, 1989.
- *Fundamentals of Speech Recognition*, L.Rabiner, B.H. Juang, Prentice Hall Signal Processing Series, 93
- *Reconnaissance automatique de la parole*, JP Haton, C. Cerisera, D. Fohr, Y. Laprie, K. Smaili, Ed. Dunod, 2006
- *Synthèse de la Parole à partir du texte*, C. d'Alessandro, E. Tzoukermann - Hermes, collection Traitement automatique des langues, 2001
- *Pattern Recognition*, R.O. Duda, P.E. Hart, D.G. Stork, Ed John Wiley & sons, inc. , 2001
- *Apprentissage artificiel*, A. Cornuéjols, L. Miclet, Ed Eyrolles, 2002