

Reconnaissance vocale par chaîne de Markov cachée

Introduction

Ce travail de diplôme a pour but de réaliser un système de reconnaissance de la parole basé sur l'utilisation des chaînes de Markov cachées.

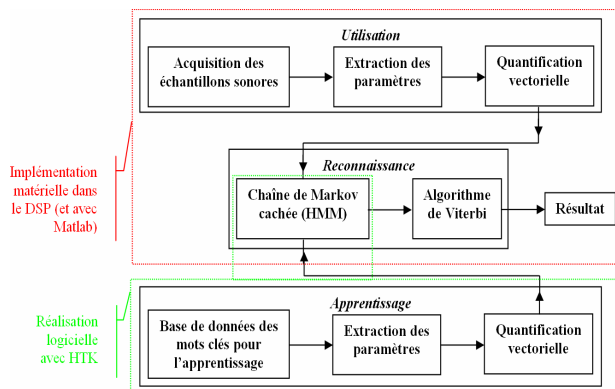
La très grande variabilité de l'aspect temporel d'un mot même lorsqu'il est répété par un même locuteur a emmené les chercheurs à utiliser des méthodes statistiques très évoluées afin de détecter et classifier des phonèmes. Les méthodes de reconnaissance basées sur les chaînes de Markov dites cachées sont au cœur des systèmes de reconnaissance les plus robustes.

Deux phases distinctes

En reconnaissance de la parole, deux phases sont essentielles :

1. une phase d'**apprentissage**, de manière à définir, modéliser et paramétrer les modèles de Markov cachés adaptés aux mots que l'on veut reconnaître.
2. Une phase de **reconnaissance**, afin de sélectionner le modèle de mot le plus vraisemblable en regard d'un mot prononcé.

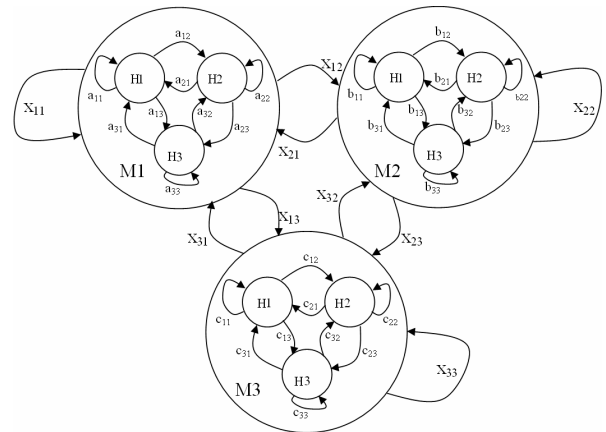
Schéma bloc du système complet :



Les chaînes de Markov cachées (HMM)

Une chaîne de Markov cachée est une machine d'états particulière dans laquelle la suite des états la plus probable n'est connue qu'à posteriori après que l'on ait analysé la vraisemblance de l'observation d'un mot que l'on considère produit par la chaîne elle-même.

Modèle de HMM à trois états :



Comme le débit de parole n'est pas constant (un même mot ainsi que ses phonèmes peuvent se prononcer plus ou moins rapidement), il est primordial de pouvoir aligner temporellement le signal. Cet alignement est réalisé grâce aux chaînes de Markov qui réalisent une segmentation optimale du signal vocal reçu.

Dès que l'on capte le signal de parole avec un micro, il est échantillonné à 16 [kHz] par un convertisseur analogique digital (CAD), puis il est traité par le DSP qui contient l'algorithme de reconnaissance.



Auteur: Sébastien DUPERTUIS
Prof. responsable: Hervé DEDIEU
Sujet proposé par: Sébastien DUPERTUIS

Hes·SO

Haute Ecole Spécialisée
de Suisse occidentale