

## Codage spatio-temporel de la parole

### Introduction

La reconnaissance vocale est au centre de l'interaction homme – machine. Ce domaine technique attire de nombreux chercheurs depuis le milieu du 20<sup>ème</sup> siècle.

La grande majorité des systèmes de reconnaissance vocale actuels sont des systèmes statistiques complexes basés sur les chaînes de Markov cachées (HMM).

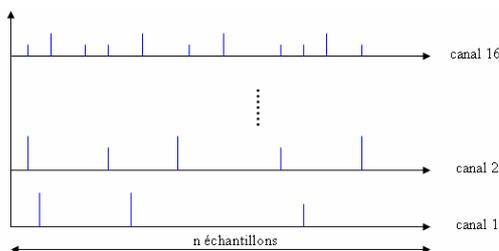
L'idée de ce travail est de s'écarter de cette méthode et de se baser sur le fait que notre système auditif, via la cochlée et l'organe de Corti, procède à une analyse spatio-temporelle du signal parole.

Le but étant de faire apparaître des invariants pour un même son voisé quel que soit le locuteur.

### Méthodologie adoptée

Ce travail de diplôme s'apparentant sensiblement à un travail de recherche pure, il consiste entièrement en un ensemble de simulations effectuées avec le logiciel Matlab.

La partie principale du travail a été de former des vecteurs acoustiques à partir d'une méthode de codage basée sur du zero-crossing et des impulsions, et de mettre en oeuvre des réseaux de neurones artificiels (ANN) afin de modéliser des systèmes capables d'identifier des voyelles voisées données.

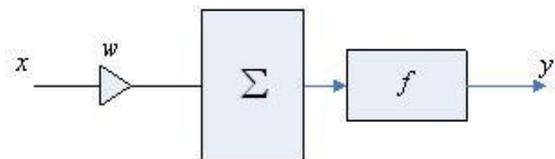


Représentation d'un échantillon sonore après codage

### Conception des systèmes

Un réseau de neurones artificiels consiste en une simplification d'un réseau de neurones biologiques; il s'agit d'un algorithme bio inspiré.

Ces réseaux possèdent de nombreux paramètres modifiés à chaque itération de l'algorithme (algorithme adaptatif) en fonction de ses vecteurs d'entrée. La ou les sorties du système permettent de prendre une décision quant à la nature du vecteur d'entrée. Cela doit permettre, dans notre cas, de classifier les voyelles voisées.



Représentation d'un neurone artificiel simple

### Entraînements et tests des systèmes

Pour que les réseaux puissent être capables d'identifier des voyelles, il faut d'abord passer par une phase d'apprentissage au cours de laquelle les paramètres sont ajustés, puis une phase de tests pour évaluer les performances du système.

### Résultats obtenus et perspectives

Plusieurs variantes d'ANN ont été testées et il en ressort que la très forte variabilité du signal parole empêche une identification correcte des vecteurs acoustiques. Les dernières simulations indiquent qu'une spécialisation des systèmes dans l'identification de sous-groupes de voyelles serait probablement judicieuse.

Il serait intéressant d'utiliser les vecteurs acoustiques élaborés dans ce travail dans des systèmes actuels à base de HMM comme alternative aux vecteurs habituellement utilisés.

Auteur: Laurent Dormond  
Répondant externe: Dr. Jean Hennebert  
Prof. responsable: Hervé Dedieu  
Sujet proposé par: Institut IICT