

Introduction production de son / parole organes : poumons, trachée, lèvres, langue ... ton de la vibration

le son est une vibration de l'air, qui est fourni par les poumons et les cordes vocales ainsi que la langue donne à l'

Traitement automatique de la parole regroupe le codage, la synthèse, la reconnaissance de la parole,

la reconnaissance de locuteurs et les émotions

les différents champs scientifiques associés sont le traitement du signal, les probabilités,

l'apprentissage artificiel, la linguistique, la reconnaissance de formes, les graphes, les transducteurs...

→ La difficulté dans ce domaine est la variabilité : - variabilité inter/intra locuteur

- conditions acoustiques (capteurs, bruits...)

- grammaire, vocabulaire...

la représentation du signal se fait grâce à l'analyse cepstrale, où on découpe le signal en parts égales.

il ressort ce que l'on appelle des cepstres (transformation d'un signal grâce à la transformée de Fourier). Le

triangle vocalique va permettre de reconnaître les phonèmes.

Outils mathématiques utilisés dans ce domaine : - espace acoustique : calcul des densités de probabilités

avec les modèles gaussiens

- Modèles de Markov et modèles de Markov cachés

- Algorithme de Viterbi

la reconnaissance de la parole est liée à ces 3 parties : Apprentissage, Ressources et Transcription, qui représentent l'architecture générale de la transcription automatique de la parole.

La partie Apprentissage nécessite différentes sources comme des textes, des enregistrements audio, des transcriptions manuelles d'enregistrements. Les Ressources représentent les modèles des langages, le dictionnaire phonétique et le

modèle acoustique. À partir d'un signal de la parole, la partie Transcription va générer soit la meilleure

hypothèse, soit les graphes de mots ou soit le réseau de confusion.

Formule statistique : $\hat{w} = \arg \max_w \underbrace{p(y|w)}_{\text{modèle acoustique}} \underbrace{P[w]}_{\text{modèle langage}}$ utilisation de logarithmes

Selon l'unité voulue, on peut avoir soit un modèle par mot si on choisit une unité inférieure à 100 mots soit une décomposition en suite d'unités élémentaires si l'unité est supérieure à 100 mots.