

BE RFIV

**VEYSSEIRE Daniel
&
BROSSEAU Fabien**

Partie I : Programmation Dynamique

Explication de la méthode :

Le système de reconnaissance en mots isolés a pour but de mettre en correspondance un signal acoustique inconnu O avec un dictionnaire de références $\{R_1, \dots, R_n\}$. La reconnaissance est achevée en retenant le mot du dictionnaire le plus proche du signal inconnu au sens du critère de distance D choisie. (Nous utiliserons une distance euclidienne sur les coefficients cepstraux)

L'algorithme d'alignement temporel dynamique (ou programmation dynamique) calcule le chemin permettant d'aligner temporellement deux séquences numériques distinctes ainsi que la distance séparant les deux séquences.

Soit $O = \{O_1, \dots, O_k\}$ et $R = \{r_1, \dots, r_p\}$ deux séquences numériques de longueurs différentes. Dans un premier temps la matrice des distances locales entre les deux vecteurs est calculée pour chaque couple (o_i, r_j) . Dans la pratique la distance euclidienne est utilisée.

A partir de la matrice de distances locales, l'algorithme de programmation dynamique établit le parcours minimisant la distance totale entre les vecteurs O et R selon le critère d'optimisation suivant :

$$c_{ij} = \min \begin{pmatrix} c_{i,j-1} + d_{i,j} \\ c_{i-1,j-1} + 2 * d_{i,j} \\ c_{i-1,j} + d_{i,j} \end{pmatrix}$$

C'est à dire que la minimisation du chemin total peut être décomposée en succession de minimisations locales.

Protocole expérimental

Pour le protocole expérimental nous allons utiliser un locuteur de référence parmi les 10 du dossier Sons OK. Puis nous allons comparer les mots prononcés par les neufs autres à ce locuteur, plus les sons que j'ai enregistré dans sonsveysseire.

Le locuteur de référence choisi est Dominique (il semble donner de bons résultats).

En utilisant la fonction rceps on obtient un taux de reconnaissance de 44,84 %.

ci dessous la matrice de confusion associée :

La case $\{i,j\}$ signifie : le mot de la ligne i a été reconnu x fois comme le mot de la colonne j .

[]	'cinq'	'cinquante'	'deux'	'diviser'	'dix'	'douze'	'egal'	'et'	'fois'	'huit'	'moins'	'multiplier'	'neuf'	'onze'	'par'
'cinq'	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'cinquante'	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0
'deux'	0	0	4	0	0	4	1	0	0	0	0	0	0	0	0
'diviser'	0	0	0	4	0	0	0	0	0	2	0	3	0	0	0
'dix'	0	0	0	0	2	0	0	0	0	1	0	0	0	0	0
'douze'	0	0	3	1	0	5	0	0	0	0	0	1	0	0	0
'egal'	0	0	0	0	0	1	4	0	0	1	0	0	0	0	0
'et'	0	0	0	0	0	0	1	4	0	3	0	0	0	0	0
'fois'	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0
'huit'	0	0	0	0	0	0	0	0	0	9	0	1	0	0	0
'moins'	0	0	0	0	0	0	0	0	0	0	6	0	0	1	1
'multiplier'	0	1	0	0	0	0	0	0	0	2	0	7	0	0	0
'neuf'	0	0	2	0	0	1	1	0	0	1	0	0	3	0	0
'onze'	0	0	0	0	0	0	0	0	0	0	0	0	0	5	0
'par'	0	0	0	0	0	0	0	0	1	0	2	0	0	0	2
'plus'	0	0	1	0	0	2	0	0	0	0	0	3	0	0	0
'quarante'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'quatorze'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'quatre'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'quinze'	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
'seize'	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
'sept'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'six'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'soixante'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'sur'	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
'treize'	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
'trente'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'trois'	0	0	0	0	0	0	0	0	0	0	0	0	0	1	2
'un'	1	0	0	0	0	0	1	0	0	0	2	0	0	0	1
'vingt'	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1
'zero'	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0

'plus'	'quarante'	'quatorze'	'quatre'	'quinze'	'seize'	'sept'	'six'	'soixante'	'sur'	'treize'	'trente'	'trois'	'un'	'vingt'	'zero'
0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	3
0	1	0	0	0	0	0	0	2	0	0	0	0	0	0	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	2	3	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
0	1	1	0	0	0	0	0	0	0	0	3	3	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
0	0	1	0	0	0	1	0	0	0	0	0	1	0	0	2
0	0	0	0	0	0	0	0	0	0	0	0	2	0	2	1
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	7	1	0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	6	0	0	0	0	0	0	0	0	0	0	0	0	4
0	0	1	5	1	0	2	0	0	0	0	0	0	0	1	0
0	0	0	1	8	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	4	0	1	0	1	0	0	0	0	2
0	0	0	0	0	0	5	1	1	0	0	0	0	0	0	3
0	0	0	0	0	0	0	5	2	0	0	0	0	0	0	3
0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	3
0	0	0	0	1	0	1	0	0	3	0	0	0	0	0	4
0	0	0	1	0	0	2	0	1	0	3	1	0	0	0	1
0	1	1	0	0	0	0	0	1	0	0	6	1	0	0	0
0	0	0	1	0	0	0	0	0	0	0	1	5	0	0	0
0	0	0	0	2	0	0	0	0	0	0	1	0	0	1	1
0	0	0	1	1	0	0	0	0	0	0	0	2	0	1	1
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	8

On remarque que les mots « dix, fois,neuf,par,seize,sur,treize,un et vingt sont peu reconnus.

Les mots « huit, multiplier,quarante,quinze,soixante,zero » sont souvent reconnus.

En utilisant la fonction melcepst de la toolbox voicebox que j'ai présenté dans le BE précédent, on obtient un taux de reconnaissance de 49,03%.

Voici la matrice de confusion associé (voir le fichier 'matrice_confusionMelcepst.jpg' pour plus de lisibilité):

[]	'cinq'	'cinquante'	'deux'	'diviser'	'dix'	'douze'	'egal'	'et'	'fois'	'huit'	'moins'	'multiplier'	'neuf'	'onze'	'par'
'cinq'	2	0	0	0	0	0	1	0	0	0	0	0	0	0	0
'cinquante'	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0
'deux'	0	0	5	0	0	3	2	0	0	0	0	0	0	0	0
'diviser'	0	0	0	5	0	0	0	0	0	1	0	3	0	0	0
'dix'	0	0	0	0	5	0	0	0	0	1	0	0	0	0	0
'douze'	0	0	2	1	0	7	0	0	0	0	0	0	0	0	0
'egal'	0	0	0	0	0	1	8	0	0	0	0	0	0	0	0
'et'	0	0	0	1	0	1	1	2	0	4	0	0	0	0	0
'fois'	0	0	0	0	0	0	0	0	5	0	0	0	0	0	0
'huit'	0	0	0	0	0	0	1	0	0	8	0	0	0	0	0
'moins'	0	0	0	0	0	1	1	0	0	0	5	0	0	0	0
'multiplier'	0	0	0	1	0	0	0	0	0	0	0	9	0	0	0
'neuf'	0	0	1	0	0	0	0	0	0	0	0	0	4	0	0
'onze'	0	0	0	0	0	0	0	0	0	0	0	0	0	7	0
'par'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	5
'plus'	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0
'quarante'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
'quatorze'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'quatre'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
'quinze'	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
'seize'	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0
'sept'	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
'six'	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
'soixante'	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
'sur'	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
'treize'	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0
'trente'	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1
'trois'	0	0	0	0	0	0	0	0	1	0	0	0	0	1	3
'un'	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0
'vingt'	0	0	0	0	0	0	1	0	0	0	3	0	0	0	0
'zero'	0	0	0	0	0	0	2	0	0	0	1	0	0	0	0
'plus'															
0	1	0	0	1	0	1	0	1	0	0	0	0	0	0	3
0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	4
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
0	0	2	0	0	0	0	0	0	0	0	3	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	2	0	0	0	0	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	2
0	1	0	0	2	0	0	0	0	0	0	0	0	0	1	1
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	7	0	0	1	0	0	0	0	0	0	0	0	0	0	1
0	1	4	0	0	0	0	0	1	0	0	0	0	0	0	4
0	0	1	1	7	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	9	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	6	0	0	0	1	0	0	0	0	1
0	0	0	0	1	0	4	0	0	0	0	0	0	0	0	4
1	0	0	0	0	0	0	6	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	7	0	0	0	0	0	0	3
0	0	0	0	2	0	0	0	0	5	0	0	0	0	0	2
0	0	0	0	0	0	2	0	0	0	4	1	0	0	0	0
0	0	0	0	0	0	0	0	1	0	0	6	0	0	0	1
0	0	0	0	0	0	0	0	0	0	0	2	3	0	0	0
0	1	0	0	3	0	0	0	0	0	0	0	0	0	4	0
0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	7

On remarque que les mots « cinq,et,quatre,seize,un et vingt » sont très faiblement reconnus, seize et un ne sont jamais bien reconnus.

En revanche les mots très fortement reconnus sont « zero, soixante, quinze, quarante, plus, onze, multiplier, huit, egal, douze »

Partie II : Modèles de Markov Cachés

Contact

Si il y a des erreurs, des remarques, des ajouts à faire, etc.

Veuillez en faire part à une de ces adresses :

wedg@hotmail.fr (Veysseire Daniel)

mickaelfabre@free.fr (Fabre Mickael)