

TP5/6

IAN

Reconnaissance de la parole par DTW

Détails sur l'implémentation

J' ai utilisé un ensemble (structure cell en Matlab), pour mieux visualiser l'ensemble des chemins (ens_chem). C'est une sorte de tableur de taille [taille du mot de Reference * taille du mot à tester].

Chaque cellule de ce tableur contient une liste:

- si c'est une liste vide c'est qu'on ne peut pas accéder à cette case. Seul la première ligne et la première colonne de l'ensemble de cellule devrait contenir une liste vide
- si un 1 est dans la liste, c'est qu'on peut accéder à cette case par la case de gauche
- si un 2 est dans la liste, c'est qu'on peut accéder à cette case par la case diagonale haut-gauche
- si un 3 est dans la liste, c'est qu'on peut accéder à cette case par la case du haut

On obtient, par exemple, pour la comparaison de alpha et Mike, le tableur ci-dessous que j'ai juste copié-collé dans un tableur excel :

Pourquoi on est tous censé avoir le même chemin minimum ?

Eh bien en observant les résultats on remarque qu'on a pas trop à se casser la tête pour trouver le meilleur chemin. En effet, il y en a un seul! Effectivement, on remarque que toutes les cellules du tableur ne sont composées que d'un seul élément (mis à part la 1ere ligne et la 1ere colonne)! Il n'y a donc qu'une possibilité pour accéder à chaque case de l'ensemble. La coloration rouge que j'ai rajouté sur le tableur est le chemin à partir de la case en bas à droite. Il n'y a qu'une possibilité pour le tracer.

Cela est du à l'algorithme que nous avons utilisé, car on compare pleins de vecteurs de taille 480 dont la probabilité pour que deux soient égaux est très faible. En calculant la distance entre deux vecteurs on est donc quasiment certain de ne jamais avoir le même résultat.

En appliquant l'algo,

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

En remplaçant les équations par a,b,c on a $c(i,j) = \min(a,b,c)$

Il y a une très petite probabilité, quasi nulle, pour qu'on ait l'un des cas suivant :

$$\min(a,b,c) = a = b$$

$$\min(a,b,c) = a = c$$

$$\min(a,b,c) = b = c$$

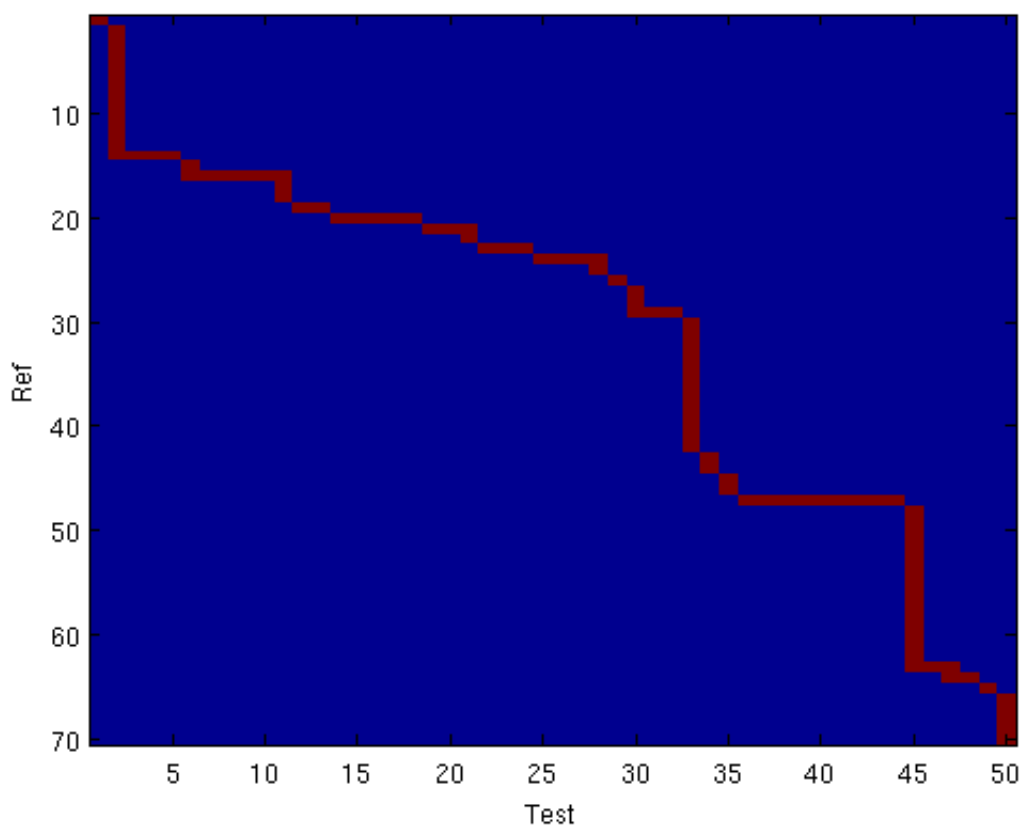
,les valeurs $c(i,j)$ étant elles même calculé à partir de distances de vecteurs.

D'ailleurs j'ai vérifié si le cas se présentait: non. Il n'y a donc qu'un chemin possible.

On est donc tous censé tous avoir les mêmes résultats et les mêmes chemins, sans même avoir besoin de prioriser la diagonale.

Même chemins, même coups, même reconnaissances pour tout le monde normalement.

Ci dessous la figure du chemin optimal obtenu, toujours avec mike comme test et alpha comme référence :



J'ai obtenu le même chemin que celui tracé sur la feuille Excel.

Reconnaissance finale

Voilà les reconnaissances obtenues :

mot alpha reconnu comme lima avec distance de 1.017

mot bravo reconnu comme bravo avec distance de 0.96232

mot charlie reconnu comme charlie avec distance de 1.0559

mot delta reconnu comme lima avec distance de 1.0076

mot echo reconnu comme echo avec distance de 1.0077

mot foxtrott reconnu comme foxtrott avec distance de 1.1446

mot golf reconnu comme golf avec distance de 0.96325

mot hotel reconnu comme hotel avec distance de 0.96672

mot india reconnu comme india avec distance de 0.95424

mot juliett reconnu comme juliett avec distance de 1.0476

mot kilo reconnu comme kilo avec distance de 0.95371

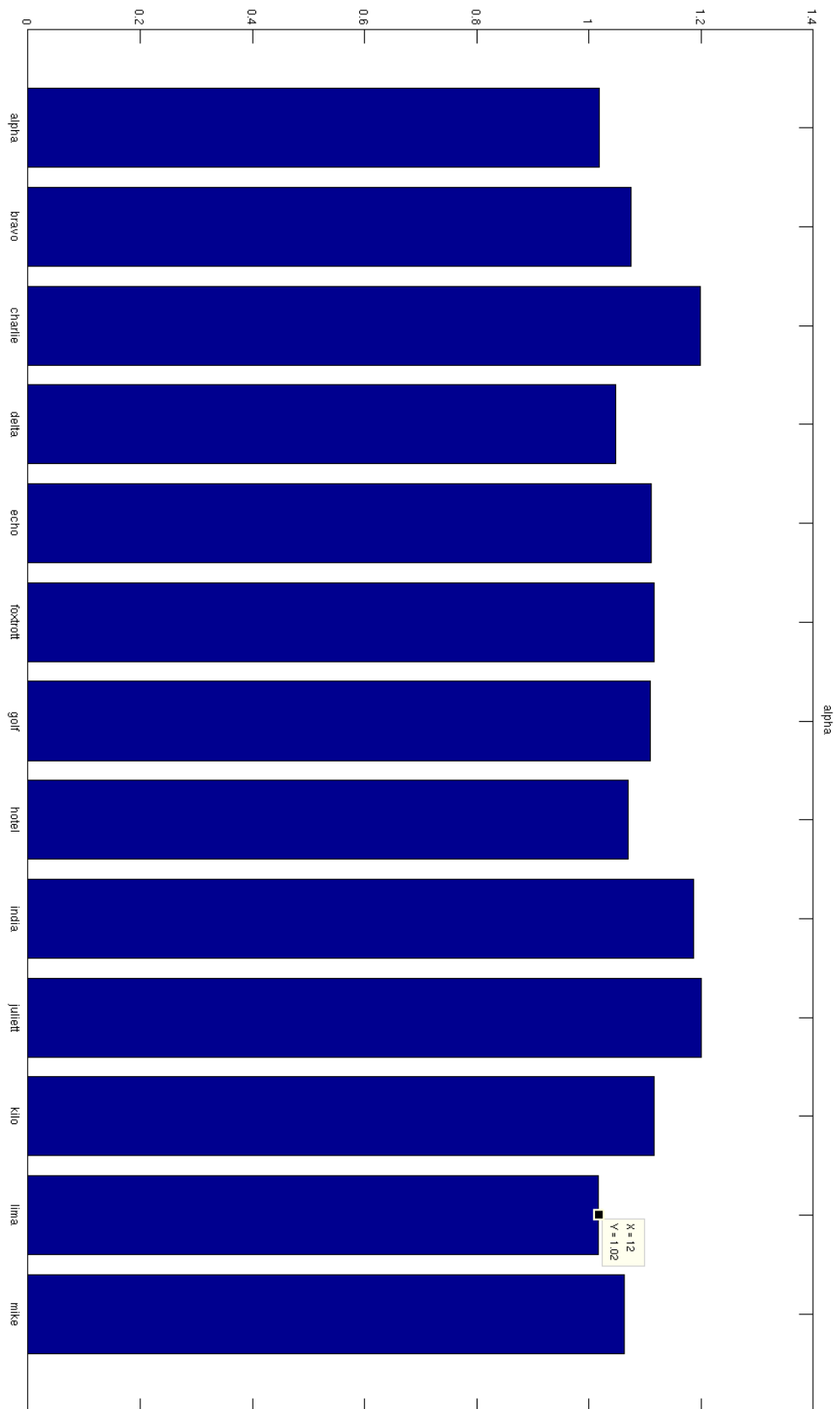
mot lima reconnu comme lima avec distance de 0.92581

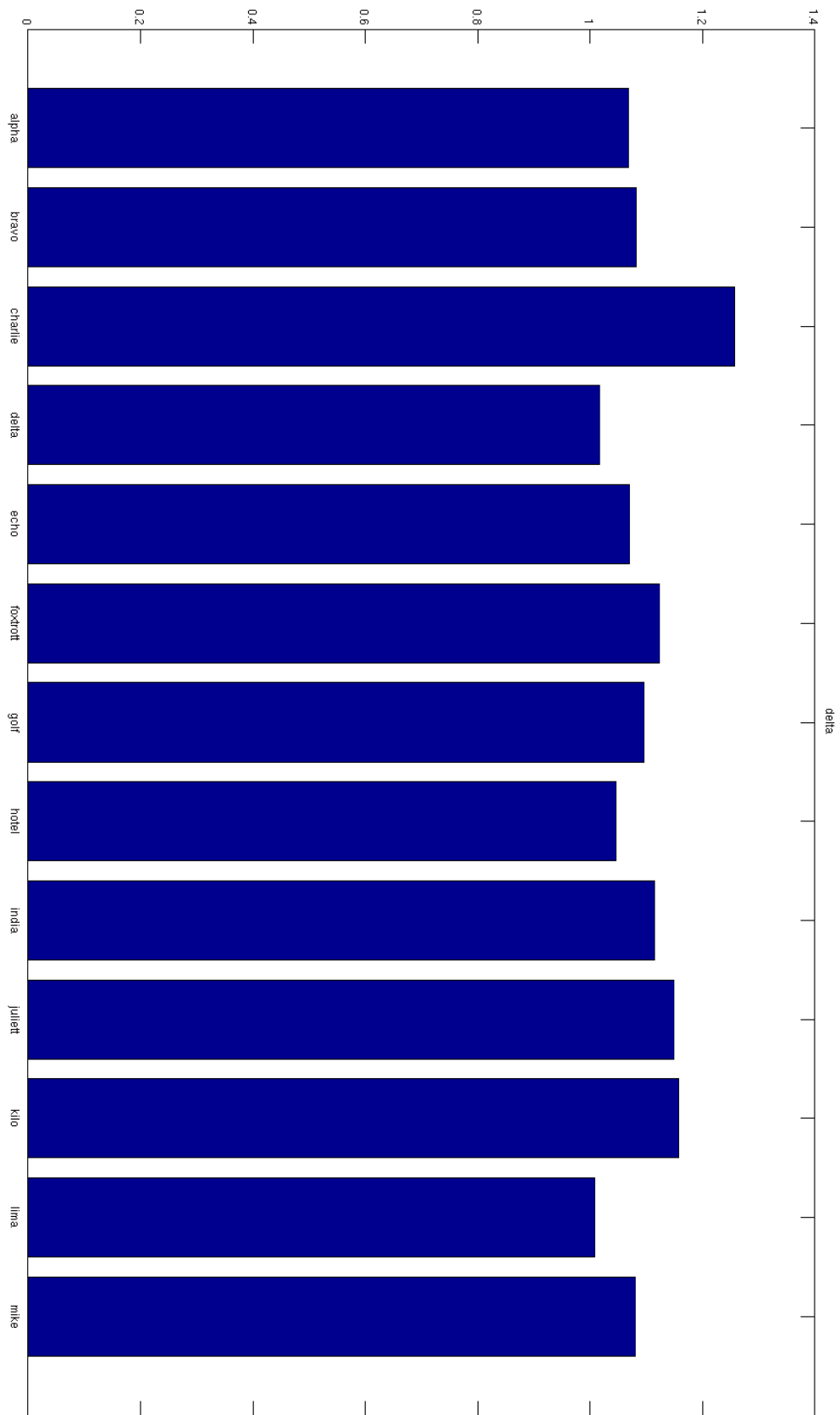
mot mike reconnu comme mike avec distance de 0.93533

Tous les mots ont été bien reconnus sauf alpha et delta qui ont été reconnus comme lima.

Ci dessous l'histogramme des distances pour alpha (voir fichier : 'histo1.png' pour plus de lisibilité)

Suivit de celles pour delta (voir fichier : 'histo2.png' pour plus de lisibilité)





On observe que les distances de alpha à alpha et de alpha à lima sont très proches. (même si le max n' est qu' à 1.2009)
(1.0187 pour alpha contre 1.0170 pour lima).

On observe que les distances de delta à delta et de delta à lima sont très proches. (même si le max n' est qu' à 1.2565)
(1.0169 pour delta contre 1.0076 pour lima).

Conclusion

Le système de reconnaissance simple apporte déjà d'assez bons résultats, mais on pourrait le rendre encore plus fiable avec d'autres paramètres que le Cepstre.

Contact

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

Veuillez m'en faire part à cette adresse :

wedg@hotmail.fr