

**S.Yahiouche ; M.Sellami**  
Département informatique  
Université Badji Mokhtar, Annaba, Algérie  
sali@swaida.com

## ***Modélisation de la lecture humaine par un système de reconnaissance de mots arabes manuscrits***

### **RÉSUMÉ.**

Notre travail s'intègre dans les efforts continus de recherche dans le domaine de la reconnaissance de l'écriture arabe manuscrite. Nous envisageons une méthode inspirée des études en psychologie cognitive ; études concernant les méthodes qu'entreprend l'être humain en lisant. Ces méthodes suggèrent que l'œil humain aperçoit les mots entiers à première vue et en saisit le sens en une fraction de secondes, s'il bute sur un mot, c'est le contexte qui l'aidera à mieux comprendre ; le contexte peut être les lettres voisines ou les mots voisins.

Malgré ces observations, moult recherches dans le domaine que nous entreprenons insistent sur le fait de granuler le mot en lettres pour ensuite reconnaître tout le mot, une opération coûteuse qui plus est, le contraire du processus naturel de lecture.

Nous essayons donc dans le présent travail de modéliser le comportement humain pendant la lecture et construire un système qui matérialise ce modèle.

**MOTS-CLÉS** Modèles perceptuels, reconnaissance de mots arabes manuscrits, effet visuels, transformée de Hough, réseaux transparents

### **1. 1. Introduction**

Les domaines qui ont suscité l'intérêt des OCR (Optical Character Recognition) depuis longtemps sont les adresses postales et les chèques, mais vu l'expansion d'Internet et de la téléphonie mobile, des cartes de crédits et de cartes bancaires, on se demande si vraiment Ces deux domaines doyens resteront de taille... En fait le monde change et le besoin de rapidité se fait de plus en plus pressant, nous commençons déjà à voir circuler les PDA (Personal Digital Agendas) où le stylo a remplacé le clavier, nous estimons aussi qu'il y aura une grande révolution sur Internet si l'on pouvait chercher de l'information sur des documents originaux manuscrits... et cela nous ramène à penser aux systèmes qui ne sont pas obligatoirement destinés à la translation intégrales des textes manuscrits en code ASCII correspondant, mais seulement créer des systèmes qui répondent à des requêtes d'utilisateurs écrites en code ASCII et qui ont pour but la recherche de certains mots clés dans le texte manuscrit.

Le domaine par excellence entamant cette voie est le domaine de reconnaissance de textes manuscrit anciens, et la méthode globale que nous entreprenons est très adaptée à ce genre d'application où la qualité dégradée du document ne supporte pas les procédures de segmentation.

### **2. Effets et Modèles Perceptuelles:**

Depuis les années soixante, les études psycho-cognitives se sont intensifiées sur les mécanismes qu'entreprend l'être humain pendant le processus de la lecture, et plus spécialement, les études et les expériences se sont multipliées pour prouver des variables linguistiques et tester leur impact sur la perception des mots. Parmi les variables les plus rencontrées dans la littérature, on peut citer :

*L'effet de supériorité du mot* ou « Word superiority effect », *L'effet de voisinage* : ou « neighbourhood effect », *L'effet de fréquence et de répétition*, *L'effet de régularité* : « Regularity effect », *L'Effet de migration des lettres*,... Beaucoup de modèles perceptuels et de systèmes de reconnaissance ont expérimentés ces effets et argumentent une grande partie de leur succès par rapport à l'explication de ces variables par leurs systèmes/modèles. On trouvera dans le travail de Grainger & Jacobs (Jacobs 1994) une étude comparative et étalée sur ses modèles.

Le modèle d'activation interactif (MAI) est effectivement le modèle perceptuel le plus utilisé dans beaucoup de systèmes de reconnaissances réalisés récemment (Pinales 2001)(Côté 1997). Il a une architecture connexionniste à trois couches (primitives, lettres, mots). C'est le modèle qui explique le mieux l'effet de supériorité du mot. Au niveau lettres, chacune est représentée par 16 primitives (segments de droites) avec une orientation particulière.

Ce modèle utilise un réseau de neurones transparent, les seuils d'activation ont été calculés comme mentionnés dans l'article original de MacClelland (McClelland 1981). Le réseau a été entraîné sur des mots

anglais imprimés de 4 caractères seulement, mais beaucoup de systèmes basés sur ce modèle ont traités des mots anglais manuscrits (Pinales 2001)(Côté 1997).

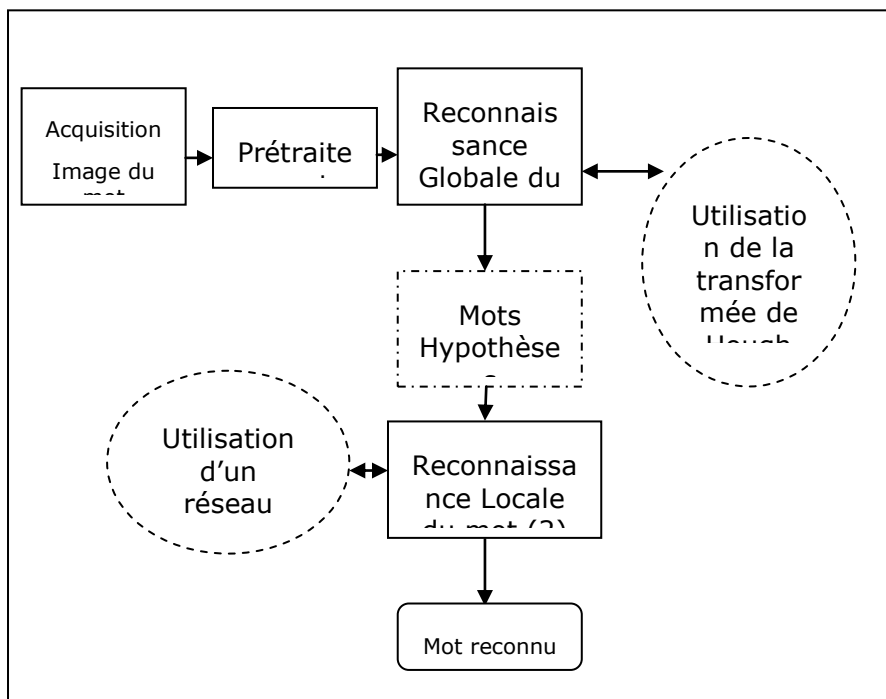
Bien que non utilisés dans des systèmes réels, les modèles perceptuels de Morton (1969) et Forster (1979) ont marqué la recherche en psycholinguistique de par leur approche similaire au comportement humain. On trouve dans (Pinales 2001),(Côté 1997) un riche aperçu sur ces modèles.

### 3. Modèle proposé :

Beaucoup de systèmes s'appuyant sur les résultats de recherche en psychologie cognitive se sont essentiellement basés sur le Modèle d'Activation Interactive (MAI). Dans ces systèmes on introduit les primitives dans une couche connectée aux lettres correspondantes et celle-ci connectée à la couche mots (voir § précédent) ; i.e ces systèmes appellent un processus globale, une démarche pratiquement ascendante qu'on trouve dans les méthodes analytiques classiques, la seule différence avec ces dernières vient du fait de l'utilisation du parallélisme, ( on ne reconnaît pas le texte de droite à gauche comme pour les HMM, mais en montant et descendant). Hors le concept fondamental qui émane des recherches en psychologie de la lecture est que pendant la lecture c'est le mot en sa globalité qui est reconnu en premier, ensuite le contexte va aider à discriminer les lettres incompréhensibles. Sortant avec ces constatations, nous avons proposé un modèle qui commence par une reconnaissance globale du mot en entrée et ce, en utilisant la transformée de Hough pour tirer les caractéristiques pertinentes de courbure et à l'aide d'un perceptron multicouche(PMC), alimentée par les cartes de la transformée de Hough, et après une phase d'apprentissage préliminaire, le PMC générera des mots hypothèses qui alimenteront la couche entrée d'un réseau de neurones transparent type (MAI) qui affinera encore plus la liste d'hypothèse émise par le PMC . Ce réseau transparent est constitué de 4 couches : mots, composantes connexes, lettres, primitives, cette dernière couche matérialise le recourt au contexte lettre en cas d'ambiguïté, se qui simule parfaitement le comportement humain.

### 4. Architecture du système :

Pour la modélisation proposée au paragraphe précédent, nous proposons l'architecture suivante :



### 5. base d'images utilisées

Nous utilisons une base de données contenant 14 400 images représentant les 48 noms des wilayas d'Algérie, divisées à parts égales pour l'apprentissage et la reconnaissance, chaque nom de Wilaya est écrit par 100 scripteur différents, un échantillon est présenté dans la figure 5 Nous comptons élargir la base aux mots contenus dans les chèques postaux (disponibles dans notre laboratoire de recherche LRI).

وہران مسکرتہ بجایہ ادرار  
متسکر بوہر داس

#### 6. Bibliographie :

- Pinales J., Reconnaissance Hors-Ligne De L'écriture Cursive Par L'utilisation De modèles Perceptifs Et Neuronaux », thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Telecom Paris, septembre 2001 ;
- Côté M. & al. Automatic Reading of Cursive Scripts Using a Reading Model and Perceptual Concepts : the PERCEPTO system , *International Journal of Document Analysis and Recognition*, Vol. 1, No. 1, pp.3 17, 1998
- Jacobs & al., Models Of Visual Word Recognition: Sampling The State Of The Art., *Journal of Experimental Psychology* , Human perception and performance, 1994, 20 (6), 1311-1334
- McClelland, J.L. and Rumelhart, D.E. An interactive activation model of context effects in letter perception : part 1 , *Psychological Review*, Vol. 88,pp. 375 407, 1981.