Cari 2004 7/10/04 11:54 Page 339

Un système multiclassifieurs appliqué au traitement de montants littéraux arabes

Nadir FARAH, Labiba SOUICI-MESLATI, Mokhtar SELLAMI

Laboratoire de recherche en informatique,
Université Badji Mokhtar Annaba Algérie
Farahnadir@Hotmail.com

RÉSUMÉ. *La détermination d'un montant littéral de chèque écrit en langue Arabe, est un problème que l'être humain résout facilement. Ce problème devient complexe lorsqu'il s'agit d'effectuer cette lecture d'une manière automatique, et offre une voie de recherche intéressante. Une approche pour la reconnaissance de montants littéraux arabes manuscrits est décrite dans cet article. La solution proposée combine diverses sources d'informations pour reconnaître les mots. L'étape de reconnaissance est effectuée par une combinaison parallèle de trois types de classifieurs (réseau neuronal, k plus proches voisins, k plus proches voisins flou) utilisant les caractéristiques globales des mots. Le contexte grammatical est utilisé pour prendre une décision finale sur les mots candidats obtenus.*

ABSTRACT. *Recognizing handwritten bank check literal amount is a problem that humans can solve easily. As a problem in automatic machine reading and interpreting, it presents an interesting field of research. An approach for recognizing the legal amount for handwritten Arabic bank check is described in this article. The proposed solution combines multiple information sources to recognize words. The recognition step is preformed with a parallel combination of three kinds of classifiers (Multilayer neural network, k nearest neighbor, fuzzy k Nearest Neighbor) using wholistic word structural features. The grammatical context is used to make final decision on obtained candidate words.*

MOTS-CLÉS : *Approche globale, multiclassifieurs, connaissances contextuelles, reconnaissance de mots arabes.*

KEYWORDS: *Holistic approach, Multiclassifier, contextual knowledge, Arabic words recognition.*

1. Introduction

La reconnaissance de l'écriture manuscrite est un domaine de recherche très actif qui a donné lieu à plusieurs travaux, notamment en ce qui concerne l'écriture latine [7, 15, 16]. La tendance actuelle des systèmes réalisés est orientée vers la combinaison de classifieurs [17] et l'intégration de sources d'informations multiples. Les systèmes réalisés se sont intéressés d'abord à un type d'écriture ou les caractères sont séparés, puis les recherches se sont accentuées sur l'écriture cursive qui est plus complexe à traiter. Dans cette voie des efforts considérables ont été déployés ces dernières années, notamment pour la reconnaissance de l'écriture Arabe [1, 3, 5].

Dans le cadre de nos recherches en reconnaissance d'écriture arabe, nous nous sommes intéressés à l'approche globale pour la reconnaissance des mots contenus dans des montants littéraux. Les caractéristiques structurelles extraites sont présentées en entrée à des classifieurs que nous avons conçus et expérimentés séparément. Ces derniers sont de type structurel [12], neuronal à représentation locale [14] et neuro-symbolique [13]. Les types d'ambiguïtés et d'erreurs produites par les différents classifieurs utilisés nous ont menés à conclure que la coopération entre plusieurs types de classifieurs et l'intégration d'informations issues de sources diverses ne peut qu'apporter un élément positif à un système de reconnaissance d'écriture manuscrite. Forts de ce constat, nous avons orienté nos recherches dans cette direction. Dans cet article, nous nous intéressons à la reconnaissance de mots Arabes manuscrits écrits en hors ligne avec un vocabulaire limité. Notre approche s'inspire de la reconnaissance humaine qui ne considère que la forme globale des mots [7]. La reconnaissance dans notre cas est effectuée par un système multiclassifieurs [9], composé de trois modules. Un analyseur syntaxique est utilisé pour valider les propositions des classifieurs.

2. Caractéristiques de l'écriture Arabe et vocabulaire considéré

La langue Arabe est très riche et difficile, de par sa structure et les possibilités qu'elle offre, elle dispose d'un alphabet de 28 caractères et s'écrit de droite à gauche. Les caractères de cet alphabet présentent des formes différentes, selon leurs positions dans le mot. Parmi les caractéristiques spécifiques des caractères nous trouvons : 10 avec un point diacritique (ز,ض,ب,ن,ظ,ذ,ف,غ,خ,ج), 03 avec deux points diacritiques (ي,ت,ق), 02 avec trois points (ش,ث), plusieurs présentent des boucles (و,ة,ه,ق,ف,م,ص …). Il existe six caractères qui imposent un espace après (ا,ز,ر,د,ذ,و), ils ne peuvent être liés que du côté droit, donc un mot peut être composé de plusieurs composantes connexes. Le vocabulaire considéré est composé de 48 mots arabes (tableau 1).

احد	تسعة	ستون	اربعمائة	ألفا	مليارات	ثمانية	خمسون
اثنان	عشر	سبعون	خمسمائة	الفان	ملايير	ثلاثمائة	الاف
ثلاثة	عشرة	ثمانون	ستمائة	مليون	سلتيم	مليارا	جزائري
اربعة	اثنا	تسعون	سبعمائة	ملايين	و	سبعة	اربعون
خمسة	عشرون	مائة	ثمانمائة	مليونا	دينار	مائتان	ألف
ستة	ثلاثون	مائتا	تسعمائة	مليونان	دنانير	مليار	سنتيمات

Tableau 1. *Vocabulaire des montants littéraux Arabes.*

3. Schéma général du système proposé

Le système proposé traite la reconnaissance de mots représentant un montant littéral écrit en arabe. Une étape d'extraction des mots du montant scannérisé a d'abord été envisagée, puis un prétraitement sur les images de mots. Ensuite une étape d'extraction de caractéristiques qui permet de déterminer les propriétés structurelles de chaque mot. A partir des caractéristiques structurelles du mot, un système multiclassifieurs, effectue la reconnaissance. Les classifieurs utilisés sont de trois types : un classifieur neuronal, un classifieur statistique de type k plus proches voisins, et un autre de type k plus proches voisins flou. Le système multiclassifieurs donne en résultat un ensemble de mots, qui seront utilisés dans une phase d'analyse syntaxique.

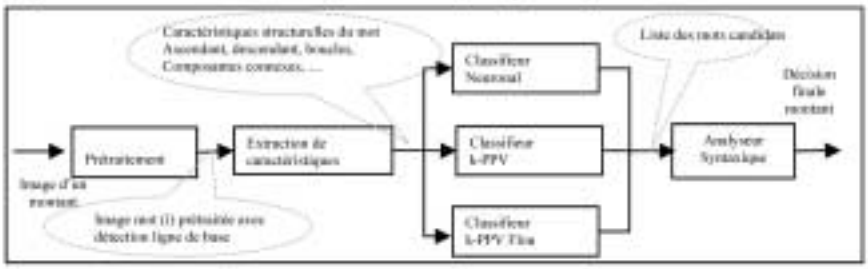

Figure 1. *Schéma général du système.*

4. Prétraitement

L'image du mot subit un ensemble de traitements avant d'en extraire les caractéristiques structurelles. Pour l'extraction des mots du montant littéral nous avons utilisé une méthode de projection verticale en plus d'une heuristique (espace entre mots

est de 1.5 fois supérieur à l'espace intra mot). L'étape de binarisation consiste à obtenir une image bimodale en utilisant une méthode de seuillage [8], elle est suivie par une étape de lissage pour diminuer les bruits [4]. L'extraction de la ligne de base est effectuée en se basant sur les projections horizontales de l'image [4].

5. Extraction des caractéristiques

Les caractéristiques structurelles utilisées dans notre approche sont les caractéristiques globales de haut niveau [7] : les descendants (D), les ascendants (A), les boucles (B), le nombre de point simple haut (PSH), le nombre de deux points hauts (PDH), le nombre de trois points hauts (PTH), le nombre de point simple bas (PSB), le nombre de deux points en bas (PDB), le nombre de composantes connexes (CC). Le tableau 2 donne les caractéristiques de quelques mots du lexique considéré.

Mot Arabe	A	D	PSH	PDH	PTH	PSB	PDB	B	CC	Mot Arabe	A	D	PSH	PDH	PTH	PSB	PDB	B	CC
			1	1				2	1		1			1		1		3	2
				2				1	1		1			2				3	2
		1			1				1		3		1					1	3
	2		1		1				2		1	1	1	1				2	2

Tableau 2. *Caractéristiques structurelles de quelques mots du lexique.*

L'extraction du contour de l'image sert à décrire l'image du mot sous forme d'une chaîne de codes de Freeman, qui représente tous les contours de l'image et sa topologie. Pour résoudre le problème de chevauchement entre parties connexes d'un mot, un algorithme de suivi de contour a été utilisé inspiré des travaux de [10]. Pour l'extraction des points diacritiques nous nous sommes basés sur les travaux de Ameur & al [2].

6. Reconnaissance structurelle du mot.

Le système multiclassifieur réalisé est composé de trois classifieurs de types différents, qui seront développés dans les sections suivantes.

Le réseau neuronal

Le réseau neuronal utilisé est de type perceptron multicouches. Ayant en entrée les caractéristiques structurelles, et donnant en sortie une classe parmi les quarante huit classes du lexique. L'apprentissage se fait par correction de l'erreur sur les neurones par la méthode de rétro propagation [6].

Les paramètres de base du système neuronal sont : Une couche d'entrée de 21 neurones correspondants aux caractéristiques structurelles selon leurs nombres d'occurrences possibles dans le lexique (voir tableau 2). Nombre de neurones en sortie: 48 neurones. Nombre de neurones couche cachée, il est calculé par une heuristique puis fixé expérimentalement à 21 neurones. La fonction d'activation est de forme sigmoïde

$$f(net_i) = \frac{1}{1+e^{-net_i}}$$

Il a été entraîné sur une base de 960 mots, qui représentent 20 exemples pour chacun des 48 mots du lexique. Nous avons obtenu un taux de reconnaissance de 96% sur cette base d'apprentissage.

Le classifieur k plus proches voisins

La conception du classifieur k-PPV débute par la création de l'ensemble d'apprentissage (base de référence), qui est constituée de M échantillons pour chacun des 48 mots du lexique, chaque échantillon est représenté par les 21 caractéristiques. Le seuil qui permettra de rejeter ou accepter les k voisins en test, est la valeur maximale sur les valeurs représentatives des distances intra-classe, la valeur représentative de distance d'une classe est la valeur maximale des distances entre les vecteurs des M échantillons de la classe pris deux par deux.

Le classifieur k plus proches voisins flou

Nous cherchons le degré d'appartenance de chaque voisin (noté Yj) par rapport aux classes références (notées classe i), pour chaque classe référence nous avons pi prototypes notés Zp, cette fonction d'appartenance [11] est de la forme (1) :

$$\mu_i(y_j) = \left[1 + \left(\max_{p=1..pi} d(y_j, Z_p) / F_d \right)^{Fe} \right]^{-1} \qquad (1)$$

Cette fonction permet d'introduire du flou Lorsque tous les degrés d'appartenance des voisins ont été testés par rapport à l'ensemble d'apprentissage, on calcule alors le degré d'appartenance de X noté : $\mu_i(X)$ par rapport à chacune des classes de ces K plus proches voisins, par la formule (2), on affecte ce mot à la classe où le degré d'appartenance est maximum.

$$\mu_i(X) = \left\{ \mu_i(y_j) * \exp(-a * d(X, y_j) / d_m) \right\} \qquad (2)$$

Où dm représente la distance euclidienne moyenne entre les mots d'une même classe dans l'ensemble d'apprentissage. a, Fe, Fd sont des constantes floues, qui ont été fixées expérimentalement aux valeurs suivantes: a=0,45, Fd=1, Fe=1. Nous utilisons un seuil d'appartenance S qui a été fixé à 0,5, ce qui a permis de rejeter les classes dont le degré d'appartenance est trop faible.

7. Résultats de la classification

Nous avons utilisé une base de 4800 mots, cette base représente 48 mots du lexique écrits par 100 scripteurs différents. 2880 mots ont été utilisés pour construire les bases de références de k-PPv et k-PPV flou, 960 mots pour l'apprentissage du classifieur neuronal et les 960 restant pour les tests. Les résultats obtenus par classifieur sont :

Pour le réseau neuronal nous avons relevé un taux de reconnaissance de 91% sur la base de test.

Pour les besoins des classifieurs k-PPV et k-PPV flou, nous avons construit quatre bases d'apprentissage (de référence), afin de nous permettre de déterminer la valeur optimale pour le paramètre k, et d'évaluer leurs taux de reconnaissance (Tableau 3).

	Taux de reconnaissance					
	k-PPV			k-PPV flou		
K	1	3	8	1	3	8
Base 1 (240 mots)	82,00	85,00	36,15	85,00	88,00	87,86
Base 2 (480 mots)	86,52	88,40	40,10	91,16	92,16	82,10
Base 3 (960 mots)	88,56	89,08	45,02	92,16	92,16	90,13
Base 4 (1200 mots)	89,08	89,08	62,00	92,16	92,16	89,47

Tableau 3. *Taux de reconnaissance de mots pour k-PPV et k-PPV flou.*

Au vu de ces résultats, nous avons remarqué que : pour k=8, les taux ont nettement baissé dans le cas du k-PPV, ceci est dû à la présence d'un nombre majoritaire d'éléments de classes distantes. Les taux de reconnaissance pour le k-PPV et le k-PPV flou sont de 89,08% et 92,16% respectivement pour k=3 valeur retenue pour notre application.

8. Post classification basée sur la syntaxe

On considérera à partir de l'étape classification qu'un mot candidat est une paire d'information, la classe du mot et sa valeur de confiance. A partir d'une grammaire utilisée par l'analyseur syntaxique, la phase post classification décide et génère un mot retenu sur l'ensemble des mots candidats.

<Milliers> ::= ألف | و + ألف | ألفان | و + ألفان | <nombre_inf_dix>+ ألاف + و | <nombre_inf_dix>+ ألاف |
<nombre_composé>+ ألف | <nombre_composé>+ ألف + و
<nombre > ::= ثلاثة | أربعة | خمسة | ستة | سبعة | ثمانية | تسعة
<nombre_compose> ::= <nombre_sup_dix>+ عشر | <nombre_sup_dix>+ عشر + و |
<nombre_sup_dix>+ و+عشر+<nombre_dizaine>

Tableau 4. *Partie de la grammaire sur les montants littéraux de chèques arabe.*

Lors de l'obtention de la liste de mots candidats par la phase reconnaissance, une sélection par vote majoritaire est effectuée sur les mots proposés. Si l'opération de vote ne peut parvenir à un résultat, on effectue une normalisation (résultats sur la même échelle) des scores obtenus puis on effectue la somme des valeurs de confiances, la valeur la plus élevée sera sélectionnée. Ensuite l'analyseur syntaxique vérifie la validité du mot obtenu par rapport aux mots déjà reconnus du montant.

9. Résultats et discussion

Un exemple d'erreur que la syntaxe a pu rectifier est donné dans le tableau 5.

Phrase correcte donnée par l'analyseur syntaxique	Phrases proposées par le système multiclassifieurs
ثلاثة آلاف و خمسون دينار	ثلاثة آلاف **و** خمسون دينار
	ثلاثة آلاف **عشر** خمسون دينار

Tableau 5. *Exemple d'erreur possible détectée par l'analyseur syntaxique.*

Le mot عشر et و ont des caractéristiques manquantes, ou mal détectées, dans le mot عشر les points diacritiques n'ont pas été pris en compte, et son premier caractère a généré une boucle, d'après la forme structurelle du mot nous avons pour les deux cas une boucle et un descendant, parmi les solutions proposées figuraient les deux mots.

Il existe des cas d'ambiguïtés syntaxiques qui ne peuvent être levés que par une information de plus haut niveau. Par exemple entre : خمسون, سبعون et تسعون .

Sur les 8% d'erreurs obtenues nous relevons, 10 % dues à une mauvaise segmentation des montants, 20 % dues à des erreurs réelles sur les mots, 30 % dues à des erreurs de classification, 40 % dues à l'absence de caractéristiques au niveau du mot manuscrit.

10. Conclusion

Dans ce travail nous avons entrepris la reconnaissance de montants manuscrits littéraux de chèques écrits en langue arabe, ce qui a impliqué plusieurs traitements : binarisation, lissage, séparation des mots du montant, identification de la catégorie du mot, analyse des mots identifiés par rapport au contexte utilisé. Le taux moyen obtenu est de 94%, nous considérons ce résultat comme très intéressant par rapport aux systèmes utilisant un seul classifieur.

L'intégration de l'analyseur syntaxique a été d'un apport très intéressant, du fait que le montant littéral du chèque n'est pas porteur que d'une information sur la structure des mots. Il convient donc d'exploiter les relations logiques, lexicales, syntaxiques et sémantiques qui existent au sein des informations extraites.

Références Bibliographiques

[1] Al Badr B., Mahmoud S. A., 'Survey and bibliography of Arabic optical text recognition', Signal processing, Vol. 41, pp 49-77, 1995.

[2] Ameur A., Romeo-Pakker K., Miled H., Cheriet M., 'Approche globale pour la reconnaissance de mots manuscrits Arabes', Actes CNED'94, 3ème Colloque National sur l'Ecrit et le Document, pp: 151-156, Juillet 1994.

[3] Amin A., 'Off-line Arabic character recognition: The state of the art', *Pattern Recognition*, vol. 31, N° 5, pp 517-530, 1998.

[4] Belaid A., Belaid Y., '*Reconnaissance des formes: Méthodes et applications*', InterEditions, 1992.

[5] Essoukhri Ben Amara, 'Sur la problématique et les orientations en reconnaissance de l'écriture arabe', CIFED 2002, pp 1-8, 2002.

[6] Jain A. K., Mao J., Mohiuddin K., 'Artificial Neural Networks: A Tutorial', *IEEE Computer, Special Issue on Neural computing*, Marsh 1996.

[7] Madhvanath S., Govindaraju V., 'The Role of Holistic Paradigms in Handwritten word Recognition', IEEE Trans. On Pattern Analysis and Machine Intelligence, vol. 23 no.2, February 2001

[8] Pavlidis T., '*Algorithms for Graphic and Image Processing*', Rockville, MD: Computer science press, 1982.

[9] Ruta D., Gabrys B., 'An Overview of Classifier Fusion Methods', *Computing and Information Systems 7(1):1-10, University of Paisley, February 2000. http://cis.paisley.ac.uk/ruta-ci0/downloads/paper1.pdf*

[10] Sari T., 'Un système de Reconnaissance de mots arabes manuscrits basé segmentation', Mémoire de Magister, Laboratoire LRI, Département Informatique, Université Badji Mokhtar Annaba, Algérie, Février 2000.

[11] Singh S., A. Amin, 'Fuzzy Recognition of Chinese Characters ', *Proc. Irish Machine Vision and Image Processing Conference (IMVIP'99), Dublin, 8-9 September, 1999.*

[12] Souici L., Aoun A., Sellami M., (2000), 'Vers une architecture multiclassifieurs pour la reconnaissance de montants de chèques arabes', Sixième conférence maghrébine en Infomatique, MCSEAI'2000, Fès, Maroc, Novembre 2000

[13] L. Souici-Meslati L., Sellami M., 'Reconnaissance de montants littéraux arabes par une approche hybride neuro-symbolique', RFIA'2002, 11th Congrès francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, Angers, Janvier 2002.

[14] Souici-Meslati L., Rahim H., Zemehri M. C Sellami M., (2002), 'Système Connexionniste à Représentation Locale pour la Reconnaissance de Montants Littéraux Arabes', CIFED'2002, Conférence Internationale Francophone sur l'Ecrit et le Document, Hammamet, Tunisie, Octobre 2002

[15] Steinherz T., Rivlin E., Intrator N., 'Off-line cursive script word recognition: A survey', International Journal on Document analysis and Recognition, IJDAR, Vol 2, pp: 90-110, 1999.

[16] Suen C.Y., 'Réflexions sur la reconnaissance de l'écriture cursive', *Actes CIFED'98, 1er Colloque International Francophone sur l'Ecrit et le Document, pp: 1-8, Québec, Canada, Mai 1998*

[17] Zouari H., Heutte L., Lecourtier Y., Alimi A., 'Un panorama des méthodes de combinaison de classifieurs en reconnaissance de formes', RFIA2002, 11th Congrès francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, pp : 499-508, Angers, Janvier 2002.