

Quelques techniques pour l'amélioration du pouvoir discriminant de primitives discrètes

Frédéric Grandidier¹ – Robert Sabourin^{2,3} – Ching Y. Suen³

¹ Centre de Recherche en Automatique de Nancy - Université Henri Poincaré, Nancy 1
11 rue de l'Université, F-88100 Saint-Dié-des-Vosges, France

² Laboratoire d'Imagerie de Vision et d'Intelligence Artificielle - Ecole de Technologie Supérieure
1100 rue Notre Dame Ouest, Montréal, H3C 1K3, Canada

³ Center for Pattern Recognition and Machine Intelligence - Concordia University
1455 de Maisonneuve Blvd West, Montréal, H3G 1M8, Canada

frederic.grandidier@iutsd.uhp-nancy.fr Robert.Sabourin@etsmtl.ca suen@cenparmi.concordia.ca

Résumé : *Cet article décrit plusieurs méthodes permettant l'amélioration du pouvoir discriminant d'ensembles de primitives discrètes. En tirant partie des possibilités offertes par la modélisation markovienne des caractères, nous proposons une technique permettant l'application de l'algorithme LDA sur nos données. Différentes méthodes de segmentation de la zone d'intérêt lors de l'extraction des caractéristiques sont présentées. Finalement une stratégie de pondération de ces différentes zones est définie. Elle permet d'intégrer l'information concernant le style d'écriture de l'échantillon traité. Les résultats expérimentaux obtenus montrent l'intérêt des différentes stratégies.*

Mots-clés : Extraction de caractéristiques, intégration d'informations, analyse discriminante linéaire, modèles de Markov cachés, reconnaissance de l'écriture manuscrite.

1 Introduction

Une des spécificités de la reconnaissance de l'écriture manuscrite est la grande variabilité intra-classe associée aux données. Cette dernière est à imputer principalement aux scripteurs, qui possèdent chacun un style propre. Une classification des différents styles rencontrés a été proposée par Tappert *et al.* [TAP 90]. Elle comporte cinq modalités différentes. Les trois dernières, correspondant à une écriture sans contrainte, sont les plus fréquemment rencontrées dans le cadre d'applications industrielles, telles que le tri automatique du courrier. Les systèmes de reconnaissance dédiés à cette tâche sont confrontés à un vocabulaire de grande taille, contraignant le concepteur à une modélisation de l'écriture au niveau caractère et de ce fait, à une phase de segmentation des mots. Cette étape du processus est complexe et ne dispose pas d'une solution absolue [CAS 95]. La sur-segmentation en graphèmes est généralement préférée. La recombinaison en caractères pouvant être effectuée par programmation dynamique.

Au niveau de la représentation des graphèmes, leur forme individuelle n'est pas toujours informative pour l'identification des caractères. La recombinaison des graphèmes en caractères peut cependant être favorisée par l'extraction de primitives discriminantes. Cette étape de bas niveau revêt une

importance particulière du fait qu'elle conditionne le reste du processus de reconnaissance.

Par cette communication, nous désirons présenter plusieurs techniques permettant l'amélioration du pouvoir de discrimination de primitives discrètes. Nos motivations sont exposées dans la section suivante. La section 3 permettra de décrire les différentes méthodes développées. Afin de valider ces approches une série d'expériences a été réalisée, elle sera décrite dans la section 4. Nos conclusions et perspectives seront exposées dans la dernière section.

2 Description de la problématique

L'objectif principal de nos travaux est l'amélioration des performances du système de reconnaissance de l'écriture manuscrite hors-ligne décrit dans [ELY 99]. Ce dernier utilise une segmentation explicite et une modélisation markovienne discrète. La combinaison par le produit cartésien de deux ensembles de primitives conduit à utiliser un total de 378 primitives (27×14). L'évaluation de ce système est réalisée à l'aide de trois corpus (voir Tab. 1) provenant d'une base de données composées d'images d'enveloppes du courrier français. Les échantillons sont classés suivant trois styles d'écriture (voir Fig. 1) : "BÂTON" (échantillons composés uniquement de caractères majuscules), "Cursif" (composés de caractères minuscules mais pouvant posséder des majuscules en début de mot) et "MiXTe" (tous les autres échantillons).

Les taux de reconnaissance obtenus sur le corpus de test pour un lexique de taille 1 000 (obtenu par tirage aléatoire dans notre lexique global contenant environ 37 000 entrées) sont présentés à la dernière colonne de Tab. 1. Étant donnés les taux de reconnaissance obtenus, nous pouvons constater que les échantillons cursifs sont moins bien reconnus que ceux des autres styles. En effet, les échantillons bâtons bénéficient d'une certaine régularité associée aux formes des lettres majuscules. Par opposition, l'écriture cursive est généralement réalisée avec plus de vélocité, conduisant à une plus grande ambiguïté des formes. De plus, l'étape de segmentation n'étant pas optimale, le caractère discret de l'écriture bâton favorise également la reconnaissance de ce style.

Corpus	Apprentissage		Validation		Test		Performance
Échantillons BÂTON	4 839	40,2%	1 174	33,8%	1 684	36%	93,4%
Échantillons Cursif	6 104	50,8%	1 831	52,7%	2 440	52,2%	85,4%
Échantillons MiXTe	1 079	9%	470	13,5%	550	11,8%	95,6%
Total	12 022		3 475		4 674		89,5%

TAB. 1 – Composition de nos corpus de données (nombre d'échantillons et pourcentage associé) et performances de notre système (taux de reconnaissance), en fonction des différents styles d'écriture.

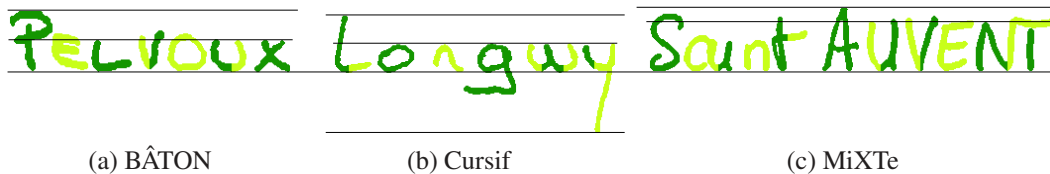


FIG. 1 – Échantillons segmentés en graphèmes et divisés par les zones d'écriture, pour les trois styles d'écriture pris en compte.

3 Différentes techniques pour l'amélioration de primitives discrètes

3.1 Prise en compte de la classe de modélisation

La tâche d'un système de reconnaissance consiste à assigner un sens ou une classe à un objet, en fonction d'un certain nombre d'observations. Dans le cadre d'un apprentissage supervisé, le concepteur dispose d'un corpus étiqueté, c'est-à-dire qu'il connaît la classe associée à chacun des objets à reconnaître. La prise en compte de l'information "classe de modélisation" lors de la construction de l'espace de représentation des données doit permettre d'obtenir des primitives plus discriminantes.

L'analyse discriminante linéaire (LDA) [DUD 01] permet de prendre en compte une telle information. Cette technique statistique consiste à chercher les composantes ou directions efficaces, permettant de réaliser la discrimination des données, en fonction des modalités d'une variable qualitative. Ces composantes permettent alors d'effectuer une projection des données dans un nouvel espace (transformation linéaire de l'original) dans lequel les échantillons sont regroupés en fonction des modalités de la variable qualitative. Son application, en utilisant l'information de classe comme variable qualitative, conduira à la construction d'un espace de représentation plus discriminant.

Cependant, cette information n'est pas toujours disponible. En l'occurrence, notre base de données est étiquetée au niveau image, c'est-à-dire que nous disposons uniquement de la chaîne de caractères présente dans l'image. Pour appliquer l'algorithme LDA, l'étiquetage au niveau graphème est nécessaire. Afin de résoudre ce problème, nous proposons d'utiliser une possibilité offerte par la modélisation markovienne. Lors de l'évaluation d'un système, l'algorithme de Viterbi [RAB 89] est utilisé. Cette procédure d'optimisation permet d'obtenir l'alignement entre les graphèmes et les classes de modélisation, conduisant à la plus forte probabilité d'observation, étant donné le système de reconnaissance en-

traîné. Cet alignement, récupéré par une procédure de *backtracking*, permet alors d'étiqueter les corpus au niveau graphème. Plusieurs étapes sont cependant nécessaires :

1. **Construction d'un premier système**
2. **Étiquetage des graphèmes** par l'intermédiaire de l'étape de *Backtracking* de l'algorithme de Viterbi
3. **Application de l'algorithme LDA** résultant en une matrice de projection
4. **Projection des échantillons** par l'intermédiaire de la matrice de projection
5. **Construction du système amélioré**

L'utilisation d'un premier système de reconnaissance ne garantit pas un étiquetage absolument exact de notre base. Cependant elle est une alternative acceptable à l'étiquetage manuel de l'ensemble de la base de données. Une vérification a permis de confirmer que le premier système retourne l'étiquette correcte pour une majorité de graphèmes.

3.2 Méthodes de segmentation de la zone d'intérêt

L'intégration d'informations discriminantes, lors de l'extraction de caractéristiques, peut être réalisée en divisant la fenêtre d'analyse en plusieurs zones et en extrayant un vecteur caractéristique distinct pour chacune d'elles [SUE 94]. Cette technique, principalement utilisée en reconnaissance de caractères isolés [TRI 96], permet une analyse locale des caractéristiques de l'écriture. Son utilisation dans le cadre du traitement de phrases conduit à la prise en compte d'informations contextuelles globales.

Dans ce contexte, nous proposons une étude concernant la méthode de segmentation de la fenêtre d'intérêt. Une première division de cette dernière sera effectuée en fonction des zones d'écriture. Il en existe trois (voir Fig. 1) :

- la zone supérieure (*Sup*) contenant les dépassements hauts,
- la zone inférieure (*Inf*) contenant les dépassements bas,
- la zone médiane (*Med*) contenant le corps du mot.

	Global	BÂTON	Cursif	MiXTe
Hauteur totale	75,3	57,2	89,7	73,3
Zone supérieure	31,3 (41,6%)	17,8 (31,1%)	41,6 (46,4%)	30,6 (41,7%)
Zone médiane	29,4 (39,0%)	32,2 (56,3%)	27,3 (30,4%)	29,9 (40,8%)
Zone inférieure	14,6 (19,4%)	7,2 (12,6%)	20,8 (23,2%)	12,8 (17,5%)

TAB. 2 – Hauteur moyenne en pixels des différentes zones d’écriture et pourcentage de la hauteur totale correspondant.

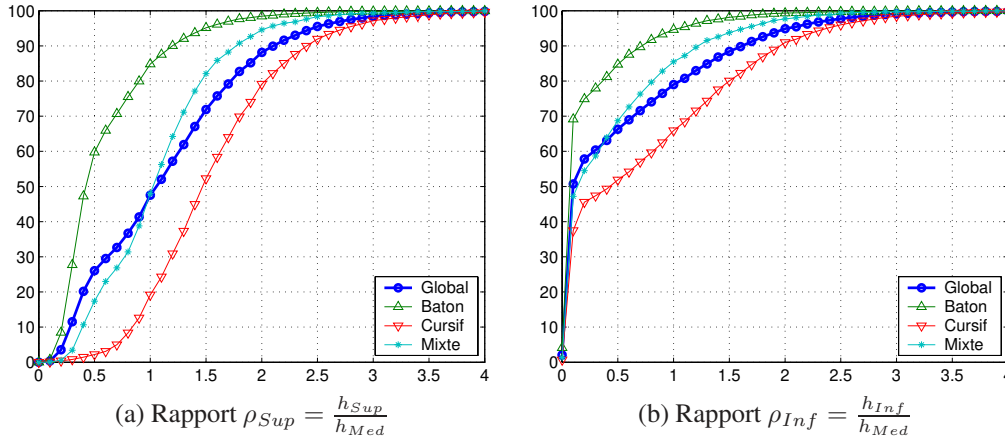


FIG. 2 – Distributions cumulées des rapports de hauteur entre les zones de dépassements et la zone médiane.

Deux droites sont utilisées pour leur définition : la ligne de base inférieure et la ligne de base supérieure. Elles sont obtenues au cours de la phase de pré-traitements [GRA 03]. Une analyse du contour inférieur permet d’obtenir les minima locaux. Après une étape de filtrage, basée sur des heuristiques, une normalisation permet de positionner les extremums restant sur une droite horizontale. La ligne de base supérieure est obtenue de manière similaire en analysant le contour supérieur. Le style d’écriture des échantillons étant inconnu lors des étapes de pré-traitements, la détection de ces zones est réalisée pour tout échantillon.

Notre système n’effectuant pas de normalisation en taille des images, la division de la zone d’intérêt par les zones d’écriture permet l’extraction d’informations semblables qu’elle que soit la taille de l’échantillon traité.

Parmi les trois zones d’écriture, la zone médiane est sans conteste celle qui contient le plus d’informations. La recherche de caractéristiques discriminantes peut être réalisée en la subdivisant. Nous proposons de tester plusieurs configurations, présentées Fig. 3. Ces dernières vont permettre de prendre en compte l’information de manière différente, particulièrement en ce qui concerne la ligature des graphèmes.

3.3 Prise en compte du style d’écriture durant l’extraction de caractéristiques

Les différentes méthodes de division de la fenêtre d’analyse proposées reposent sur la détection des zones d’écriture. L’aspect heuristique de sa mise en œuvre conduit à l’apparition d’artefacts, comme la présence de dépassements pour des échantillons bâtons (provenant de la différence de hauteur des caractères d’un même mot ou d’effets de style du scripteur, voir Fig. 1(a)). Afin d’évaluer quantitativement ces

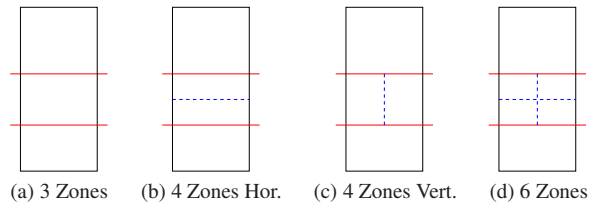


FIG. 3 – Illustration des différentes méthodes de division de la fenêtre d’analyse.

phénomènes, les hauteurs des différentes zones d’écriture ont été estimées en nombre de pixels, pour l’ensemble de nos échantillons (voir Tab. 2).

L’analyse de ces données montre l’apparition de zones de dépassement de taille non-négligeable pour les échantillons bâtons. Afin de compléter notre étude, les deux rapports de hauteur suivants ont été définis :

$$\rho_{Sup} = \frac{h_{Sup}}{h_{Med}} \quad \text{et} \quad \rho_{Inf} = \frac{h_{Inf}}{h_{Med}} \quad (1)$$

où les quantités h_i sont les hauteurs des trois zones d’écriture. Les distributions cumulées de ces rapports sont présentées Fig. 2. Lors de l’extraction de primitives, l’heuristique permettant de considérer un dépassement haut significatif correspond à une valeur de $\rho_{Sup} > 1/3$. Le graphique associé à ce rapport permet de quantifier que 60% des échantillons bâtons ont un dépassement haut significatif (contre 99% pour les échantillons cursifs). Pour l’ensemble de ces exemples, des primitives caractérisant l’écriture cursive vont être générées et donc conduire à l’introduction de confusions lors de la reconnaissance. Concernant les dépassements bas,

l'heuristique associée est : $\rho_{Inf} > 1/2$, c'est-à-dire que 15% des exemples bâtons et environ 30% des exemples mixtes ont des dépassements bas significatifs.

L'utilisation des zones d'écriture lors de l'extraction de caractéristiques conduira à une meilleure prise en compte de l'information pour les échantillons cursifs, mais induira une certaine confusion pour les échantillons bâtons. La connaissance *a priori* du style des échantillons serait une information intéressante. Cependant son obtention est difficile, voire impossible. Nous proposons alors une alternative : réduire l'influence des zones de dépassements en définissant une stratégie de pondération.

Soit $\mathbf{x}_i = \{x_j^i\}$ le vecteur caractéristique associé à un échantillon dans un espace à N dimensions, avec $x_j^i \in [0, 1]$. Soit $\mathbf{z}_k = \{z_j^k\}$ un centre de gravité résultant d'une quantification vectorielle de cet espace :

$$z_j^k = \frac{1}{N_{V_k}} \cdot \sum_{i=1}^{N_{V_k}} x_j^i \quad (2)$$

où N_{V_k} est le nombre d'échantillons rattachés à \mathbf{z}_k . Soit n_{Sup} , n_{Med} et n_{Inf} , le nombre de dimensions de l'espace de représentation relatif aux trois zones ($N = n_{Sup} + n_{Med} + n_{Inf}$). Soit $\mathbf{p}_i = \{p_j^i\}$ le vecteur de pondération associé à l'échantillon \mathbf{x}_i , nous définissons ses composantes de la manière suivante :

$$\begin{cases} p_j^i = \min(1, \rho_{Sup}^i), & 0 < j \leq n_{Sup}, \\ p_j^i = 1, & n_{Sup} < j \leq n_{Sup} + n_{Med}, \\ p_j^i = \min(1, \rho_{Inf}^i), & n_{Sup} + n_{Med} < j \leq N. \end{cases} \quad (3)$$

La prise en compte de ce vecteur de pondération est réalisée directement lors de l'étape de quantification vectorielle, en utilisant une distance euclidienne pondérée :

$$d_{Ep}(\mathbf{x}_i, \mathbf{z}_k) = \sqrt{\frac{1}{P_i} \sum_{j=1}^N p_j^i (x_j^i - z_j^k)^2} \quad (4)$$

avec

$$P_i = \sum_{j=1}^N p_j^i = (\rho_{Sup}^i \times n_{Sup}) + n_{Med} + (\rho_{Inf}^i \times n_{Inf}) \leq N$$

Cette approche (Pond1) a l'avantage de ne pas modifier les valeurs des vecteurs caractéristiques \mathbf{x}_i . Par contre elle ne peut plus être utilisée après LDA puisque chaque nouvelle composante du vecteur caractéristique est une combinaison linéaire de celles de \mathbf{x}_i . La seule alternative afin de prendre en compte la pondération des zones après LDA consiste à créer un nouveau vecteur $\mathbf{y}_i = \{y_j^i\}$ tel que $y_j^i = p_j^i \times x_j^i$. Cette seconde approche (Pond2) revient à modifier directement les valeurs des vecteurs caractéristiques.

4 Expérimentation

4.1 Protocole

Différentes techniques permettant l'amélioration d'ensembles de primitives discrètes ont été présentées. Deux espaces de représentations numériques ont été développés afin

d'évaluer leur pertinence. Le premier (CCV) permet de quantifier les concavités présentes dans les graphèmes. Sa mise en œuvre est similaire à celle exposée dans [HEU 94]. Seules les configurations de concavités continues comportant plus de deux sondes sont conservées, le taux de pixels d'écriture analysés est également pris en compte. Le vecteur caractéristique contient alors 45 composantes.

Le second espace de représentation (DDD) est celui proposé par Oh et Suen [OH 98]. Il est basé sur l'extraction des distributions de distances directionnelles. Notre mise en œuvre [GRA 03] diffère du fait que les images utilisées ne sont pas normalisées en taille. Par contre, les distances mesurées sont normalisées en fonction de la taille de la zone d'analyse. L'extraction de ces caractéristiques sur des graphèmes rend inutile l'étape de *tiling* proposée par Oh et Suen. Finalement, chaque zone d'analyse est divisée en quatre cellules de taille identique, ce qui conduit à un vecteur caractéristique de taille 64.

La modélisation discrète utilisée par notre système impose une étape de quantification vectorielle. L'algorithme LBG [LIN 80] est utilisé pour sa rapidité d'exécution. Des ensembles de primitives contenant 256 primitives ont été construits. Cette taille conduit généralement aux meilleurs taux de reconnaissance, dans le cadre de notre application. Les résultats présentés (en pourcentage) sont tous obtenus lors de l'utilisation d'un lexique de taille 1 000. L'évaluation du choix de segmentation de la zone d'intérêt ainsi que l'influence de l'utilisation de l'algorithme LDA sont présentées Tab. 3. Une extraction de primitives sur le rectangle englobant des graphèmes a été réalisée afin de servir de référence pour l'évaluation des stratégies de division de la zone d'intérêt. L'influence de la stratégie de pondération est évaluée par l'intermédiaire de Tab. 4.

4.2 Analyse des résultats

Concernant l'intégration de l'information de classe par LDA, nous pouvons remarquer qu'elle conduit globalement à un gain de performances. Lors de l'extraction de primitives sur les rectangles englobants (cf. lignes 1-2 de Tab. 3), l'amélioration est présente quel que soit l'espace de représentation utilisé et quel que soit le style d'écriture. Nous pouvons conclure que cette méthode doit être utilisée par les systèmes de reconnaissance basés sur des modèles de Markov cachés discrets effectuant une extraction de caractéristiques sur les rectangles englobants. En mode reconnaissance, l'utilisation de cette approche n'introduit qu'un temps de calcul supplémentaire négligeable, relatif à la projection de l'échantillon dans le nouvel espace de représentation.

Concernant les espaces de représentation utilisés, nous pouvons remarquer que l'espace DDD conduit à de moins bonnes performances avant LDA que l'espace CCV. Par contre la tendance s'inverse une fois l'analyse discriminante réalisée. Dans [OH 98] l'auteur annonce un grand pouvoir discriminant de la technique d'extraction proposée. Cependant il utilise des réseaux de neurones comme classificateur. Cette approche permet d'effectuer la discrimination nécessaire par l'intermédiaire des couches cachées. Dans notre cas, l'analyse discriminante doit être réalisée afin d'obtenir des

Espace	Division	Avant LDA				Après LDA			
		Global	BÂTON	Cursif	MiXTe	Global	BÂTON	Cursif	MiXTe
CCV	Rectangle	88,3	94,1	82,7	95,1	89,0	95,2	83,1	95,6
DDD	Englobant	83,6	92,1	75,7	92,7	92,7	96,5	89,1	97,4
CCV	3 Zones	89,6	92,6	86,4	94,7	89,9	94,2	85,7	95,8
DDD		86,3	88,8	83,0	93,3	93,2	95,6	90,5	97,6
CCV	4 Zones	90,9	94,2	87,3	96,9	90,8	95,2	86,6	95,4
DDD	Hor.	87,5	90,7	83,6	94,9	93,6	95,9	91,3	96,7
CCV	4 Zones	89,9	93,1	86,2	96,0	90,4	95,2	85,9	95,3
DDD	Vert.	88,7	92,2	85,0	93,8	93,8	96,8	90,9	97,3
CCV	6 Zones	90,9	94,4	87,3	96,2	91,0	95,2	86,7	96,7
DDD		88,5	93,5	83,5	95,3	92,9	96,1	89,8	96,5

TAB. 3 – Évaluation des stratégies d’intégration de l’information de classe et de division de la zone d’analyse.

Espace + Division	Pondération	Avant LDA				Après LDA			
		Global	BÂTON	Cursif	MiXTe	Global	BÂTON	Cursif	MiXTe
CCV 3 Zones	Sans	89,6	92,6	86,4	94,7	89,9	94,2	85,7	95,8
	Pond1	89,5	92,8	85,8	96,0	89,7	94,2	85,3	95,6
	Pond2	89,7	93,2	86,2	94,9	90,0	94,5	85,9	94,7
DDD 3Zones	Sans	86,3	88,8	83,0	93,3	93,2	95,6	90,5	97,6
	Pond1	87,1	89,8	83,9	93,1	93,3	95,8	90,7	97,1
	Pond2	86,1	89,3	82,3	93,3	93,1	95,5	90,7	96,7

TAB. 4 – Évaluation des deux stratégies de pondération du vecteur caractéristique.

performances similaires. Elle permet de prendre en compte la richesse de la représentation de l’information.

La comparaison des taux de reconnaissance de Tab. 3 avec ceux du système standard (cf. Tab. 1), permet de remarquer que les différents ensembles de primitives construits au cours de cette phase expérimentale conduisent à de meilleures performances. De plus, il est important de remarquer que ce gain s’accompagne d’une diminution de la complexité du système de reconnaissance, puisque le nombre de primitives est réduit de 378 pour le système standard à 256.

La mise en évidence de l’influence de la stratégie de division de la zone d’intérêt se fait par l’intermédiaire de Tab. 3. Nous pouvons constater que l’utilisation de cette technique conduit bien au résultat escompté : une amélioration des taux de reconnaissance associés aux échantillons cursifs. Cependant, ce gain s’accompagne d’une dégradation des performances associées aux échantillons bâtons, comme nous l’avions pressenti. Le gain étant plus important que la perte et le nombre d’échantillons cursifs plus important que celui des bâtons (cf. Tab. 1), la segmentation de la zone d’intérêt conduit globalement à un gain de performances. Les remarques ci-dessus sont valables pour les deux espaces de représentation évalués.

L’augmentation du nombre de zones conduit également à une amélioration des taux de reconnaissance. La meilleure performance est obtenue lors de l’utilisation de six zones pour l’espace de représentation CCV. Par contre, pour l’espace DDD, elle est obtenue pour quatre zones avec division horizontale de la zone médiane. En effet, la combinaison “Six Zones-DDD” conduit à un trop grand nombre de dimensions

de l’espace de représentation utilisé et se traduit par une réduction de la fiabilité du résultat de quantification vectorielle. En conclusion, le choix de la stratégie de division de la zone d’intérêt lors de l’extraction de primitives doit être effectué en fonction de l’espace de représentation utilisé et en limitant sa dimension.

Afin de compenser la perte de performance associée aux échantillons bâtons lorsque l’extraction de caractéristiques est réalisée à l’aide de plusieurs zones, deux stratégies de pondération ont été proposées. L’analyse des colonnes étiquetées “Global” de Tab. 4 permet de constater que leur utilisation n’influence pas de manière significative la performance globale des systèmes de reconnaissance. Par contre, une analyse en fonction du style d’écriture permet de remarquer que les deux stratégies de pondération conduisent bien au résultat prévu, à savoir la réduction de la perte de performances pour les échantillons bâtons. Cependant, les taux de reconnaissance atteints sont encore bien inférieurs à ceux obtenus lors de l’extraction des caractéristiques sur les rectangles englobants. Par exemple, pour les ensembles de primitives basés sur les concavités, l’extraction sur les rectangles englobants conduit à un taux de reconnaissance de 95,2% après LDA alors qu’il n’est que de 94,5% avec prise en compte de la pondération.

Concernant les échantillons cursifs, nous pouvons remarquer que la prise en compte de la pondération a un effet différent. Dans certains cas elle conduit à une légère diminution des taux de reconnaissance. La raison est que pour certains échantillons cursifs, la prise en compte d’une stratégie de pondération réduira l’influence de certains dépassements

hauts ou bas qui peuvent être significatifs et ainsi entraînera une certaine confusion dans le système de reconnaissance. Une analyse détaillée permet de remarquer que la première stratégie de pondération, celle ne modifiant pas le vecteur caractéristique, semble être favorable aux ensembles de primitives DDD alors que la seconde stratégie de pondération améliore légèrement les taux de reconnaissance associés au système utilisant les primitives de concavités. Notre interprétation de cette différence est que l'information contenue dans chaque composante du vecteur DDD est plus riche et donc plus sensible à une altération de sa valeur.

5 Conclusions et perspectives

Différentes techniques permettant l'amélioration du pouvoir discriminant d'ensembles de primitives discrètes d'un système de reconnaissance de l'écriture manuscrite ont été présentées. Dans un premier temps nous avons introduit une technique permettant d'intégrer l'information "classe de modélisation" lors de la construction d'ensembles de primitives, pour un système utilisant une modélisation markovienne des caractères. Nous avons ensuite évalué, dans un contexte de reconnaissance de phrase, l'apport de la division de la zone d'intérêt lors de l'extraction de caractéristiques. Cette technique est bien adaptée à la reconnaissance de l'écriture cursive mais introduit une certaine confusion pour l'écriture bâton. Afin de palier ce problème, nous avons introduit deux stratégies de pondération permettant de réduire l'influence des composantes du vecteur caractéristique associées aux zones de dépassements. Les différentes expériences réalisées ont montré la pertinence de ces techniques, en particulier pour la reconnaissance d'échantillons cursifs.

L'utilisation de l'algorithme LDA conduit à la construction d'un nouvel espace de représentation. Lors de cette étape, il est possible d'effectuer une réduction de dimensions en ne conservant que les composantes les plus significatives. Cette possibilité n'a pas été évaluée durant notre phase expérimentale, elle fait partie des travaux futurs. Une remarque découlant de notre étude est qu'il existe une différence significative de performances entre les styles cursif et bâton. Les techniques introduites permettent de les améliorer, mais toujours plus pour un style que pour l'autre. Trouver le moyen de coupler les meilleures performances de chaque style fait partie des perspectives de notre travail.

Remerciements

Ce travail a été supporté par le Service de Recherche Technique de La Poste à Nantes, France, par l'École de Technologie Supérieure et le Center for Pattern Recognition and Machine Intelligence à Montréal, Canada.

Références

- [CAS 95] CASEY G., LECOLINET E., Strategies in character segmentation : A survey, *3rd ICDAR*, vol. 1, Montréal, Canada, 1995, pp. 1028-1033.
- [DUD 01] DUDA R. O., HART P. E., STORK D. G., *Pattern Classification*, John Wiley & Sons, New York, USA, 2001, 2nd edition.
- [ELY 99] EL-YACOUBI A., GILLOUX M., SABOURIN R., SUEN C. Y., An HMM-based approach for off-line unconstrained handwritten word modeling and recognition, *IEEE Trans. on PAMI*, vol. 21, n° 8, 1999, pp. 752-760.
- [GRA 03] GRANDIDIER F., Un nouvel algorithme de sélection de caractéristiques – Application à la lecture automatique de l'écriture manuscrite, PhD thesis, École de technologie supérieure - Univ. du Québec, 2003, <http://www.livia.etsmtl.ca/publications/theses.htm>.
- [HEU 94] HEUTTE L., Reconnaissance de caractères manuscrits : application à la lecture automatique des chèques et des enveloppes postales, PhD thesis, Univ. de Rouen, 1994.
- [LIN 80] LINDE Y., BUZO A., GRAY R. M., An algorithm for vector quantizer design, *IEEE Trans. on Communication*, vol. 28, n° 1, 1980, pp. 84-95.
- [OH 98] OH I. S., SUEN C. Y., Distance features for neural network-based recognition of handwritten characters, *IJDAR*, vol. 1, n° 2, 1998, pp. 73-88.
- [RAB 89] RABINER L., A tutorial on hidden Markov models and selected applications in speech recognition, *Proceedings of the IEEE*, vol. 77, n° 2, 1989, pp. 257-285.
- [SUE 94] SUEN C. Y., GUO J., LI Z. C., Analysis and recognition of alphanumeric handprints by parts, *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 24, n° 4, 1994, pp. 614-631.
- [TAP 90] TAPPERT C., SUEN C. Y., WAKARA T., The state of the art in on-line handwriting recognition, *IEEE Trans. on PAMI*, vol. 12, n° 8, 1990, pp. 787-808.
- [TRI 96] TRIER O., JAIN A., TAXT T., Feature extraction methods for character recognition - A survey, *Pattern Recognition*, vol. 29, n° 4, 1996, pp. 641-662.