



HAL
open science

Modélisation de l'écrit en ligne à l'aide de réseaux bayésiens et de caractéristiques relationnelles

Rudy Sicard, Thierry Artières, Eric Petit

► To cite this version:

Rudy Sicard, Thierry Artières, Eric Petit. Modélisation de l'écrit en ligne à l'aide de réseaux bayésiens et de caractéristiques relationnelles. Colloque International Francophone sur l'Écrit et le Document (CIFED'06), Sep 2006, Fribourg, Suisse. pp.181-186. hal-00120435

HAL Id: hal-00120435

<https://hal.science/hal-00120435v1>

Submitted on 18 Dec 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Modélisation de l'écrit en ligne à l'aide de réseaux bayésiens et de caractéristiques relationnelles

Sicard Rudy^{1,2} – Artières Thierry² – Petit Eric¹

¹ France Telecom RD/TECH/IDEA
28 chemin du vieux chêne – 38000 Meylan - France

² UPMC, Laboratoire d'informatique de Paris 6
8 rue du capitaine Scott - 75015 Paris - France

{Rudy.Sicard, Eric.Petit}@orange-ft.com, Thierry.Artieres@lip6.fr

Résumé : *Nous proposons une nouvelle approche pour la modélisation de l'écriture en ligne. Celle-ci s'appuie sur l'utilisation de caractéristiques locales et relationnelles. L'utilisation de caractéristiques relationnelles permet de concevoir des modèles présentant des propriétés de robustesse vis à vis de perturbations sur l'ordre et le sens des traits, et de traits superflus. Ces modèles peuvent être utilisés dans des applications diverses comme des tâches de reconnaissance et de segmentation partielle de séquences.*

Mots-clés : segmentation , écriture en-ligne, réseaux bayésiens dynamiques.

1 Introduction

Un certain nombre de modèles ont été proposés pour traiter des données séquentielles, les reconnaître et les segmenter. En général, ces modèles posent des hypothèses d'indépendance simplificatrices facilitant leur apprentissage. C'est notamment le cas des Modèles de Markov Cachés (MMC) très utilisés dans ce domaine. Néanmoins, de nombreuses extensions des MMC ont été proposées pour la prise en compte des dépendances entre observations de la séquence, comme les MMC autorégressifs et les MMC segmentaux [ODK 96]. Dans le domaine du traitement de l'image, notamment de la segmentation, les champs de Markov aléatoires (MRF) introduisent des dépendances locales entre observations par le biais de caractéristiques relationnelles entre pixels voisins [HEL 97].

Récemment, le développement des modèles conditionnels, en particuliers les champs aléatoires conditionnels (CRF) [LCP 01], a permis de s'affranchir des principaux défauts des MMC et des modèles génératifs en général. En effet, ces modèles ne nécessitent aucune hypothèse simplificatrice d'indépendance sur la distribution des données. Ces modèles ont généralement été appliqués à des données textuelles, en exploitant des caractéristiques discrètes et se sont montrés performants pour des tâches de classification et de segmentation sur ce type de données. Ils sont cependant peu appropriés à d'autres types de

données et pour d'autres traitements de séquences qui peuvent s'avérer utiles dans le cadre du traitement de l'écriture en ligne.

L'objectif de ce travail est de développer des modèles efficaces pour le traitement de séquences pour diverses tâches telles que la reconnaissance et la segmentation de séquences graphiques, et aussi pour des tâches plus générales d'analyses de séquences, comme la reconnaissance partielle, le rejet, le diagnostic. Le diagnostic consiste à fournir une évaluation sur la qualité d'une séquence par rapport à un modèle donné.

Nous proposons ici un modèle basé sur des caractéristiques locales classiques (comme dans les MMC) mais incluant aussi des caractéristiques relationnelles. Ces caractéristiques relationnelles portent sur chaque paire d'observations, et traduisent une partie de la dépendance statistique entre observations. Cette approche conduit à un modèle génératif implémentant une distribution est proche de celle des MRF. Dans le contexte de l'écriture, les caractéristiques relationnelles correspondent à des relations spatiales entre observations graphiques. L'information spatiale permet d'améliorer les performances en reconnaissance. Cependant, cette information n'est que très peu employée pour la segmentation ([MAR 04][CKI 01]), excepté dans le cas de la reconnaissance de caractères asiatiques où des méthodes ad hoc ont été proposées [LCC 96]. Dans le cas d'une écriture bruitée ou de mauvaise qualité (e.g. écriture d'un enfant), en présence de distorsions dans l'ordre du tracé ou le sens des tracés, cette information spatiale est nécessaire pour obtenir un traitement robuste. Nous montrons alors que le modèle que nous proposons, exploitant ces caractéristiques spatiales relationnelles, répond bien à l'objectif de robustesse.

Nous présentons la stratégie de modélisation (§2) puis nous discutons des algorithmes d'inférence et d'apprentissage (§3 et §4). Enfin nous expliquons comment ces modèles peuvent être utilisés pour effectuer des tâches de reconnaissance partielle ou complète de tracés manuscrits et fournissons des résultats d'expérience sur des données d'écriture en ligne (§5).

2 Modélisation relationnelle de séquences

La probabilité $p(x|y)$ d'une séquence d'observations $x=(x_1, \dots, x_T)$ conditionnellement à la séquence d'états $y=(y_1, \dots, y_T)$ (c.a.d. la segmentation) est définie par :

$$p(x|y) = \prod_{t=1}^T p(x_t | x_1^{t-1}, y) \quad (1)$$

avec la notation $x_1^{t-1} = (x_1, \dots, x_{t-1})$.

Une telle distribution de probabilité étant difficile à estimer, il est d'usage d'introduire des simplifications comme les hypothèses d'indépendance utilisées dans les MMC. Néanmoins, ces hypothèses ne permettent pas de prendre en compte des dépendances complexes ou entre observations éloignées. Des modèles plus riches ont été proposés dans lesquels un état génère une sous séquence d'observations à l'aide d'une loi conjointe, rendant ces observations interdépendantes localement, c'est le cas de modèles de trajectoires et segmentaux [ODK 96].

Nous nous intéressons ici à l'approximation de $p(x_t | x_1^{t-1}, y_1^T)$ sous la forme d'un produit de fonctions potentielles :

$$p(x_t | x_1^{t-1}, y) \approx \frac{1}{Z(y, x_1^{t-1})} f(x_t, y_t) \prod_{i=1}^{t-1} g(x_t, x_i, y_t, y_i) \quad (2)$$

où f et g sont des fonctions potentielles quelconques et $Z(y, x_1^{t-1})$ est un facteur de normalisation garantissant que l'expression est une probabilité. La fonction f exprime les dépendances locales entre une observation x_t et sa variable d'état associée y_t , alors que la fonction g exprime la dépendance entre deux observations issues de leur état respectifs. La forme de l'Eq. 2 est très proche des "pairwise MRF" fréquemment utilisés dans le domaine de la reconnaissance et de la segmentation d'images [HEL 96]. La modélisation par paire apparaît comme une approche efficace pour estimer des lois de probabilités complexes sur un ensemble de variables et constitue un compromis intéressant entre généralité d'expression et complexité. Elle permet de prendre en considération les dépendances entre une observation x_t et l'ensemble des autres observations par le biais des liens deux à deux. Notre modèle est plus général que les "pairwise MRF" dont la fonction g a pour forme $g(y_i, y_t)$; les observations x_t n'ont plus de dépendance directe entre elles.

En utilisant les Eq. 1 et 2, la probabilité de la séquence complète peut s'écrire :

$$p(x_1^T | y) \approx \frac{1}{Z(y)} \prod_{t=1}^T f(x_t, y_t) \prod_{i=1}^{t-1} g(x_t, x_i, y_t, y_i) \quad (3)$$

La difficulté principale de l'Eq. 3 réside dans le facteur de normalisation $Z(y)$ qui peut conduire à des algorithmes coûteux. Ce terme se calcule toutefois raisonnablement dans certains cas, par exemple si toutes les fonctions potentielles sont gaussiennes. Dans ce travail, nous utilisons plus simplement des fonctions potentielles normalisées, c'est-à-dire des lois de probabilités, et le modèle de l'équation (3) s'exprime selon :

$$p(s, r | y) = \prod_{t=1}^{n^k} p(s_t / y_t) \prod_{i=1}^{t-1} p(r_{t,i} / y_t, y_i) \quad (4)$$

où s et r sont deux ensembles de caractéristiques extraites de x , s'interprétant comme une représentation dual de la séquence d'entrée x . La séquence s est formée de caractéristiques locales notées s_t , extraites de x_t . La matrice $r = \{r_{t,i}\}_{1 \leq t, i \leq T}$ exprime les liens deux à deux entre observations, où $r_{t,i}$ représente la (ou les) caractéristiques relationnelles liant x_t et x_i . Par exemple, on peut choisir que $s_t = x_t$ et $r_{t,i} = x_t - x_i$. La figure 1 illustre le modèle de l'Eq. 4 sous la forme d'un réseau Bayésien dynamique (DBN) pour une séquence de taille 4.

Divers modèles peuvent être implémentés en fonction des caractéristiques locales et relationnelles utilisées et de la distribution $p(r_{t,i} | y_t, y_i)$ retenue. Tout d'abord, afin d'étudier l'intérêt des caractéristiques locales et relationnelles (Cf. §5.3) nous utiliserons une version pondérée de la loi de probabilité de l'Eq. (4), du type :

$$\log p(s, r | y) = w \sum_{t=1}^{n^k} \log p(s_t / y_t) + (1-w) \sum_{t=1}^{n^k} \sum_{i=1}^{t-1} \log p(r_{t,i} / y_t, y_i) \quad (5)$$

Egalement, on peut construire un modèle local purement séquentiel, ou un modèle purement relationnel, ou un modèle hybride. Par exemple, un modèle purement relationnel s'obtient en définissant la relation $r_{t,i} = x_t - x_i$ avec une distribution gaussienne pour la loi $p(r_{t,i} | y_t, y_i)$. Ce modèle n'utilise pas de caractéristique locale et n'utilise pas l'ordre des observations. Dans ce cas la segmentation est guidée uniquement par les liens entre observations sans tenir compte de leur ordre. Cette propriété est importante pour le cas où le signal d'entrée (ici l'écriture) est potentiellement très déformée à la fois dans son aspect mais également dans l'ordre dans lequel sont tracés les différentes parties d'un caractère etc. Nous reviendrons sur ces aspects dans la section expérimentale.

3 Inférence et segmentation

Etant donnée une séquence d'observations x , la segmentation consiste à trouver le meilleur étiquetage y^*

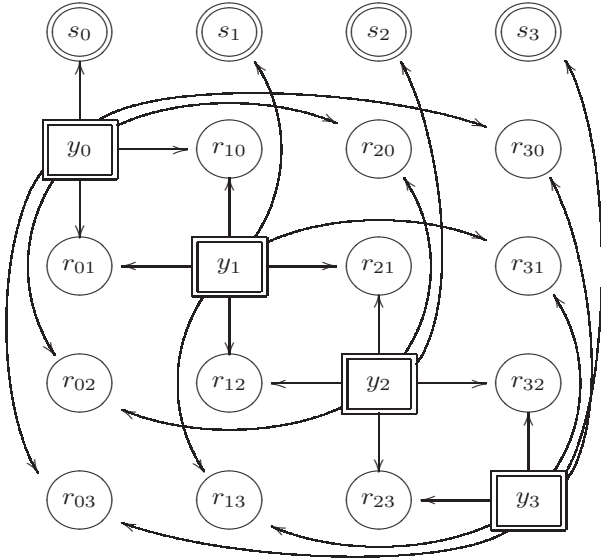


FIG. 1. Réseau Bayésien dynamique illustrant le modèle de l'équation (4) pour un signal d'entrée de longueur 4.

maximisant la probabilité *a posteriori* $P(y/x)$. Cependant, l'inférence dans le DBN donné par l'Eq. 4 (cf. Fig. 1) est un problème d'inférence NP-complet en raison des nombreuses boucles du réseau Bayésien. C'est pourquoi, nous proposons d'approximer $P(y/x)$ par une fonction de la forme : $q(y) = \prod_{i=1} q_i(y_i)$ qui est une forme

d'approximation appropriée lorsque la masse de probabilité de la distribution est concentrée autour de la meilleure segmentation. On peut en effet montrer que le produit des probabilités marginales est la meilleure approximation de cette forme, selon la divergence de Kullback-Leibler [CAR 03]. On peut choisir d'estimer les lois marginales à l'aide de l'algorithme d'inférence Belief Propagation (BP) (que nous ne détaillerons pas ici), qui reste efficace même pour des réseaux complexes avec boucles [WFR 01]. En prenant l'estimateur MAP (Maximum A Posteriori) de cette approximation, on peut alors déterminer la segmentation optimale :

$$y^* = \arg \max_y q(y) \approx \arg \max_y p(y/x) \quad (6)$$

Cette approximation sera également utile lors de l'apprentissage du modèle.

Cet algorithme d'inférence a une complexité proportionnelle au carré de la longueur de la séquence multiplié par le carré de la taille du modèle (c.a.d. le nombre d'états).

4 Apprentissage

Considérons tout d'abord le cas totalement supervisé où la base d'apprentissage comprend des séquences étiquetées, c.a.d. incluant un ensemble de N séquences d'observations $X = \{x^1, \dots, x^N\}$ ainsi que leurs segmentations respectives $Y = \{y^1, \dots, y^N\}$. Ici, chaque segmentation y^k est une séquence d'états de même

longueur (n^k) que la séquence d'observation correspondante x^k . D'après Bayes :

$$P(\Theta|X, Y) = \frac{P(X|Y, \Theta)P(\Theta|Y)}{P(X|Y)} \quad (7)$$

avec :

$$p(X|Y, \Theta) = \prod_{k=1}^N \left[p(s^k, r^k | y^k, \Theta) \right]$$

où (s^k, r^k) est la représentation duale de x^k (§2). Alors, selon l'Eq. (3) :

$$p(X|Y, \Theta) = \prod_{i=1}^N \left[\prod_{t=1}^{n^k} p(s_t^k | y_t^k, \theta) \prod_{j=1}^{t-1} p(r_{t,j}^k | y_t^k, y_j^k, \theta) \right] \quad (8)$$

où s_t^k est la $t^{\text{ième}}$ caractéristique de la séquence s^k . En

notant α_l et $\beta_{l,m}$ les paramètres des fonctions potentielles pour les caractéristiques locales et relationnelles, on définit les termes

$$p(s_t | y_t = q_l, \theta) = p(s_t | \alpha_l) \quad \text{et}$$

$$p(r_{t,j} | y_t = q_l, y_j = q_m, \theta) = p(r_{t,j} | \beta_{l,m}).$$

En supposant de plus que $P(\Theta)$ se factorise selon :

$$p(\Theta) = \prod_{l=1}^L p(\alpha_l) \prod_{m=1}^L p(\beta_{l,m}) \quad (9)$$

Alors toute méthode d'apprentissage basée sur la probabilité *a posteriori* des paramètres peut être envisagée. On montre de plus, en combinant les équations (7) à (9), que les paramètres des fonctions potentielles peuvent être appris indépendamment. En pratique, la segmentation complète y^k d'une séquence d'apprentissage n'est pas connue, l'étiquetage d'une séquence d'apprentissage est réduit à une information de classe. De fait, le calcul de la probabilité *a posteriori* des paramètres nécessite de sommer sur toutes les segmentations Y possibles :

$$p(\Theta|X) = \sum_Y p(\Theta|X, Y) p(Y|X) \quad (10)$$

Cependant ce calcul pose plusieurs difficultés : $p(Y|X)$ n'est pas connue, la sommation sur Y est prohibitive voire non calculable, et les paramètres ne sont plus indépendants et donc ne peuvent plus être optimisés séparément. Nous proposons d'utiliser une approche EM pour contourner ces difficultés. Par cette approche, on peut montrer que les paramètres s'estiment à nouveaux indépendamment les uns des autres. En suivant la dérivation de l'algorithme EM dans [TAN 93], qui maximise l'espérance du logarithme de la probabilité *a posteriori* des paramètres, on obtient :

$$\text{Etape } E : \text{calculer } q(y^i) \approx p(x^i | y^i, \theta^t)$$

$$\text{Etape } M : \theta^{t+1} = \arg \max_{\theta} \prod_y \left[p(\theta) p(x^i | y^i, \theta) \right]^{q(y)}$$

On peut alors en déduire les formules de ré-estimations suivantes pour les paramètres β pour l'étape M :

$$\hat{\beta}_{l,m} = \arg \max_{\beta_{l,m}} \left[p(\beta_{l,m}) \prod_{i=1}^N \prod_{t=1}^{n^k} \prod_{j=1}^{t-1} p(r_{t,j}^k / \beta_{l,m}) q(y_i^k = l, y_j^k = m) \right] \quad (11)$$

où $q(y_i^k = l, y_j^k = m)$ est une probabilité qui reflète l'appartenance des observations x_i^k et x_j^k respectivement aux états l et m . Notons que ces derniers termes sont calculés par approximation avec l'algorithme d'inférence BP décrit dans la section 3.

5 Expérimentations

Comme nous l'avons précisé précédemment, divers modèles peuvent être exprimés selon ce schéma. Nous en donnons quelques exemples ci-après. Concernant la partie expérimentale, les tracés utilisés sont des chiffres manuscrits provenant de 12 scripteurs de la base de données Unipen.

5.1 Qu'apprend un modèle relationnel ?

Afin de montrer l'intérêt de la modélisation basée sur les relations spatiales et d'illustrer son fonctionnement, nous étudions ce qui a été appris par un modèle relationnel exploitant des caractéristiques relationnelles $r_{l,i} = x_l - x_i$ modélisées par des distributions gaussiennes $p(r_{l,i} | y_l, y_i)$. Le modèle est ici composé de trois états et a été appris sur les trois exemples du symbole II de la Figure 2-b. La figure 2-a schématise les distributions de probabilités $p(r_{l,j} | \beta_{l,m})$ apprises sur les caractéristiques relationnelles. La matrice 3x3 fournit une représentation graphique de chaque distribution gaussienne, une par case, correspondant à des couples d'états. Ainsi, la case située à la $l^{\text{ème}}$ ligne et la $m^{\text{ème}}$ colonne montre la valeur des paramètres $\beta_{l,m} = (\mu_{l,m}, \Sigma_{l,m})$. Deux éléments sont représentés dans chaque case. La moyenne $\mu_{l,m}$ est la translation moyenne entre deux observations dans les états respectifs l et m , elle est représentée par un segment partant du centre de la case. Centrée autour de l'extrémité du vecteur moyen, une ellipse représente la dispersion (modélisée par la matrice de covariance $\Sigma_{l,m}$) autour de la valeur moyenne. Par exemple, la représentation dans la case (1,3) montre que les observations issues du 3^{ème} état sont à droite de celles issues du 1^{er} état, avec une dispersion verticale importante et une dispersion horizontale faible. Remarquons que les translations entre observations d'un même état sont de valeur moyenne nulle et que leur dispersion donne une information sur l'orientation du tracé dans cet état (cases (1,1); (2,2); (3,3)).

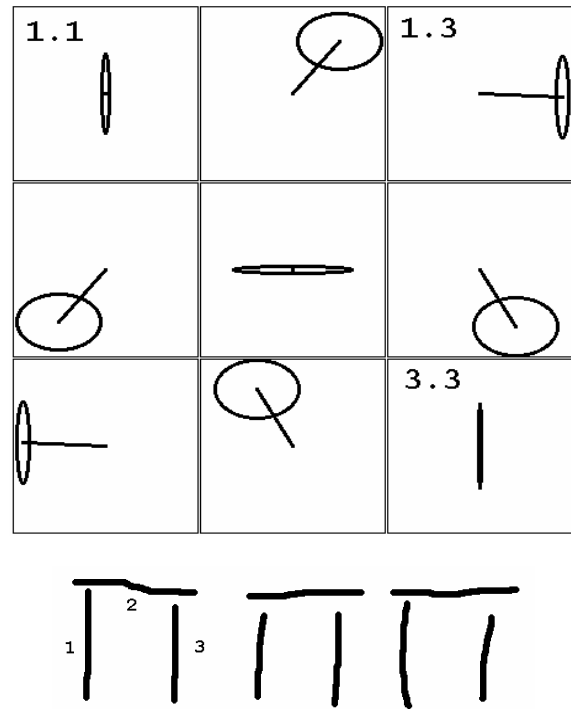


FIG. 2. Représentation schématique des distributions sur les caractéristiques relationnelles d'un modèle purement relationnel (en haut) appris avec trois exemples d'apprentissage (en bas).

5.2 Qualité de l'écriture et segmentation robuste

Evaluer la qualité d'un signal d'écriture peut être utile pour différents objectifs. Cela peut se faire globalement, comme dans les procédures de rejet pour la reconnaissance de caractères écrits où on compare la vraisemblance d'un tracé par rapport à un seuil. Il y a cependant des situations qui exigent une analyse plus précise comme dans l'analyse du tracé manuscrit chez l'enfant afin de détecter des troubles d'écriture. Par ailleurs, il y a aujourd'hui un intérêt croissant pour l'automatisation d'outils de diagnostic utilisant l'écriture et le dessin comme support ([BGC 05][HDD 87][GHP 05]). Pour de telles tâches, il est nécessaire de mettre en œuvre une analyse intelligente afin de détecter les parties mal écrites ou mal dessinées, les parties absentes ou superflues. De plus, il est nécessaire d'identifier les distorsions locales et globales, de détecter par exemple que certaines parties d'une lettre n'ont pas les bonnes proportions, qu'un *o* est écrit en sens inverse, ou encore que le point d'un *i* est trop haut. L'étape de segmentation est cruciale pour obtenir de telles informations. Elle doit être robuste et doit permettre de bien segmenter les parties bien écrites sans être influencée par l'absence de certaines parties ou l'existence de parties mal écrites. Comme nous l'avons évoqué les modèles classiques

(MMC) sont inappropriés dans ce cas. Ils sont très sensibles à l'ajout de tracés superflus et inadaptés aux changements d'ordre dans le tracé. Nous présentons maintenant quelques expériences mettant en évidence les propriétés de robustesse de notre modèle. Nous montrons que l'utilisation de caractéristiques relationnelles apporte un gain de robustesse concernant la segmentation d'un tracé bruité. La figure 3 montre, à droite, un exemple de tracé segmenté par un modèle appris sur l'exemple de gauche. Bien que le tracé de droite ait été dans un ordre différent et partiellement à contresens on constate que le modèle est capable de le segmenter correctement.

Pour évaluer plus systématiquement la robustesse à des perturbations sur l'ordre et le sens du tracé nous avons créé artificiellement une base de tracés perturbés. Un signal d'écriture original est perturbé en lui appliquant trois types de perturbations élémentaires : permutation de deux traits (un trait est ici une partie de tracé sans levée de stylo), l'inversion du sens d'écriture d'un trait et la division d'un trait en deux traits. La répétition d'un nombre de ces perturbations élémentaires introduit un niveau élevé de bruit au niveau temporel (sens, ordre, continuité, ...). Le niveau de perturbation d'un signal peut être réglé en modifiant le nombre de ces perturbations élémentaires. Nous avons construit une base de test dont les signaux sont perturbés par N perturbations élémentaires de chaque type, en faisant varier N . Pour un niveau de perturbation N , nous appliquons N "séparations", puis N "permutations" et enfin N "inversions". La figure 4 montre l'évolution de la robustesse de la segmentation en fonction du niveau de perturbation N . Plus précisément, la courbe indique le pourcentage d'observations du tracé pour lesquelles la segmentation du signal originale diffère de celle du signal perturbé. Nous constatons que le premier niveau de perturbation ($N=1$) génère un taux d'erreurs de 4%, qui augmente lentement jusqu'à atteindre 9% pour $N=10$. En considérant le niveau important de perturbation introduit pour $N=10$, ces résultats montrent la grande robustesse de ce modèle relationnel vis-à-vis de perturbations temporelles.

5.3 Reconnaissance

Notre modèle peut également être utilisé pour la reconnaissance de caractères en utilisant un modèle pour chaque classe de caractères. Cependant la section précédente a montré que le modèle peut associer une probabilité élevée à une séquence incomplète vis-à-vis du modèle (lorsque tous les états n'apparaissent pas dans la segmentation). Afin de l'utiliser en reconnaissance, il est donc nécessaire d'ajouter un mécanisme simple pour prendre en compte cette information d'incomplétude. Ce mécanisme est obtenu en estimant, durant l'apprentissage, la probabilité pour chaque état d'être observé. Il est ainsi possible d'évaluer la probabilité qu'une segmentation particulière soit correcte pour le modèle. Lors de la reconnaissance, le score de chaque classe est obtenu en multipliant la probabilité de la séquence par la probabilité que la segmentation soit correcte.

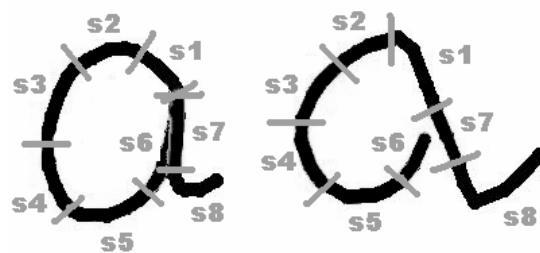


FIG. 3 - Figure centrée sur une colonne Deux exemples de "a" différents correctement segmentés par le même modèle. Les labels (s1,s2,...) indiquent la segmentation

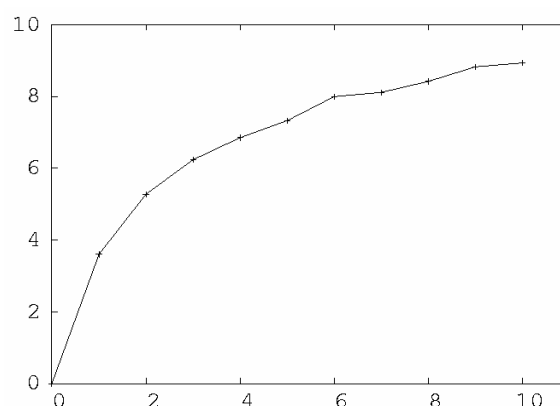


Fig. 4 - Pourcentage d'observations pour lesquelles la segmentation du signal originale diffère de celle du signal perturbé (en ordonnée) en fonction du niveau de bruit (abscisse).

Nous rapportons des résultats de deux séries d'expériences en mode monospace sur les chiffres de la base Unipen. Ces expériences ont été réalisées sur l'écriture de 12 scripteurs ayant fourni suffisamment de tracés, on utilise 20 tracés par chiffre en apprentissage et au moins 10 tracés en test. Les expériences ont été répétées 10 fois pour chaque scripteur en faisant varier les tracés d'apprentissage et de test. Les résultats fournis ici sont des taux de reconnaissance moyens sur les 10 répétitions et sur les 12 scripteurs.

Dans la première série d'expériences (Table 1), nous comparons un MMC à un système purement relationnel (§2). Les modèles ont 5 états dans les deux cas et les MMC ont une topologie classique gauche-droite. Les résultats montrent que les modèles relationnels sont significativement meilleurs que les MMC. En effet le taux d'erreur est plus faible de 30%. Dans une seconde série d'expériences (fig. 5), nous étudions l'importance relative des caractéristiques relationnelles et locales pour la reconnaissance. Nous comparons des modèles où les caractéristiques locales et relationnelles ont des poids différents. Lorsque le poids est de 1.0, seules les caractéristiques locales sont utilisées, et lorsque le poids

est de 0.0, seules les caractéristiques relationnelles sont utilisées.

TAB. 1 – Comparaison d'un modèle purement relationnel et d'un MMC pour la reconnaissance de chiffre isolés en ligne.

Système	Taux de reconnaissance
MMC	95.3%
Modèle relationnel	96.9%

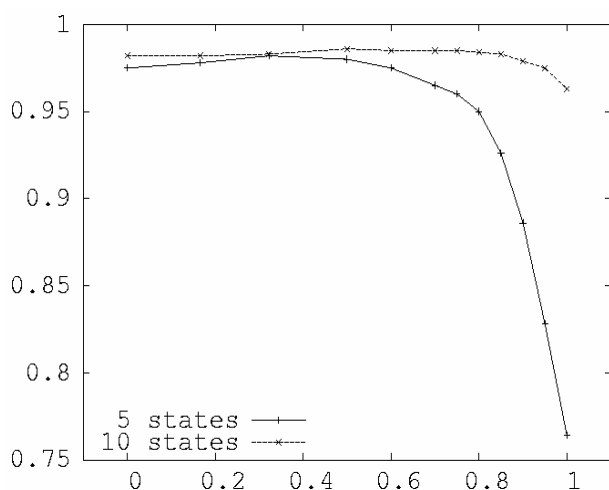


FIG. 5 – Performance en classification de chiffres de deux modèles (à 5 et 10 états) exploitant des caractéristiques relationnelles et locales, cf. Eq. (5), en fonction du poids w . Un poids égal à 0 correspond à un modèle purement relationnel, un poids égal à 1 à un modèle purement local, les valeurs intermédiaires correspondent à des modèles hybrides.

Nous fournissons des résultats pour deux tailles de modèle différentes (5 ou 10 états). Nous constatons d'abord que l'usage des caractéristiques locales seules induit un taux de reconnaissance faible, et que ce taux est très sensible au nombre d'états du modèle. Par comparaison, les modèles purement relationnels ont un taux de reconnaissance plus élevé et sont bien moins sensibles à la taille du modèle. De plus, nous observons que la combinaison des deux types de caractéristiques améliore sensiblement les performances, indépendamment du nombre d'états. En effet, selon le nombre d'états, l'ajout de caractéristiques relationnelles permet de réduire le taux d'erreur de 92% pour les modèles à 5 états, et de 62% pour les modèles à 10 états.

6 Conclusion

Nous avons présenté un nouveau schéma de modélisation pour le signal écrit en ligne. Il permet de construire divers modèles probabilistes (incluant les MMC) qui exploitent conjointement des caractéristiques locales et relationnelles. Nous avons précisé les algorithmes d'inférence et d'apprentissage pour ces modèles et montré leur intérêt pour le traitement de l'écriture en ligne comme, la segmentation partielle et

robuste, le diagnostic, la reconnaissance. De plus, nous avons montré comment ces caractéristiques relationnelles apportent plus de robustesse pour des déformations telles que l'ajout de tracé superflus, l'absence de portions du tracé, les variations d'ordre et de sens dans le tracé.

Références

- [BGC 05] Bara F., Gentaz E., Colé P., "Early handwriting acquisition and its difficulties", Reading and writing, 2005.
- [CAR 03] Cardoso J-F., "Dependence, correlation and Gaussianity in independent component analysis", Journal of Machine Learning Research, 12/2003, Vol. 4, pp 1177-1203.
- [CKI 01] Cho S.-J., Kim J.H., "Bayesian Network Modeling of Strokes and Their Relationships for On-Line Handwriting Recognition", International Conference on Document Analysis and Recognition, 2001.
- [GHP 05]. Glenat, L. Heutte, T. Paquet, D. Mellier. "Computer-Based Diagnosis of Dyspraxia: the MEDDRAW project", Conference of the Int. Graphonomics Society, IGS 2005.
- [HDD 87] Hamstra-Bletz L., DeBie J. and Den Brinker B. "Concise evaluation scale for children handwriting", L.S.Z, 1987.
- [HEL 97] K. Held and all, "Markov Random Field Segmentation of Brain MR Images". IEEE Trans. on medical imaging, Vol. 16, No. 6, 1997.
- [LCC 96] Liu, W.-K. Cham, and M.M.Y. Chang, "Stroke Order and Stroke Number Free On-Line Chinese Character Recognition Using Attributed Relational Graph Matching", International Conference on Pattern Recognition, ICPR, 1996.
- [LCP 01] Lafferty J., McCallum A., and Pereira F., Conditional random fields: "Probabilistic models for segmenting and labeling sequence data", International Conference on Machine Learning, 2001.
- [MAR 04] Marukatat S., Artières T., "Handling spatial information in on-line handwriting recognition", International Workshop on Frontiers in Handwriting Recognition, 2004.
- [ODK 96] Ostendorf M., Digalakis V.V., Kimball O.A., "From HMMs to Segment Models: A Unified View of Stochastic Modelling for Speech Recognition", IEEE Transactions of Speech and Audio Processing, Vol. 4, No. 5, pp. 360-378, 1996.
- [PEA 88] Pearl J., "Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference", San Francisco, CA, Morgan Kauffmann Ed., 1988.
- [TAN 93] Tanner M., "Tools for statistical inference", Springer Verlag, NY, third edition.
- [WFR 01] Weiss Y. and Freeman W., "Correctness of belief propagation in graphical models of arbitrary topology", Neural Computation, Vol. 13, pp. 2173-2200, (2001).