

Utilisation de champs de Markov en approches globale et analytique pour la reconnaissance de l'écriture manuscrite

Christophe Choisy

► **To cite this version:**

Christophe Choisy. Utilisation de champs de Markov en approches globale et analytique pour la reconnaissance de l'écriture manuscrite. Journées Jeunes Chercheurs, GRCE, 1999, Tours, France, 2 p, 1999. <inria-00098783>

HAL Id: inria-00098783

<https://hal.inria.fr/inria-00098783>

Submitted on 26 Sep 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.

Utilisation de champs de Markov en approches globale et analytique pour la reconnaissance de l'écriture manuscrite

Christophe Choisy
Abdel Belaid

LORIA, Campus scientifique, B.P. 235
54506 Vandoeuvre-lès-Nancy, France

1 Introduction

La reconnaissance de l'écriture utilise principalement deux approches : globale et analytique ([1, 2]). La première considère la présence de caractéristiques sur l'ensemble du mot, ce qui lui permet de bien absorber les petites variations locales, mais la limite à un vocabulaire distinct. De plus, cette vision du mot introduit une certaine complexité qui ne peut être tolérée sur un vocabulaire étendu. La seconde approche permet de s'affranchir de ces limites, mais nécessite une interprétation locale (segmentation en graphèmes, reconnaissance par regroupement) qui est plus sensible au bruit et aux distorsions locales.

Notre approche se propose de tirer avantage des deux méthodes en essayant de réduire la complexité de la méthode globale en l'appliquant sur des entités plus petites (lettres), et d'utiliser la méthode analytique pour la reconnaissance en laissant le système déterminer automatiquement les meilleures frontières. Nous avons utilisé les champs de Markov causaux (NSHP-HMM, [3]) pour la méthode globale pour ses capacités de prise en compte de la nature de l'image. La reconnaissance du mot utilise un HMM dont l'estimateur local est le NSHP-HMM.

Nous avons testé cette approche sur des mots issus des montants de chèques postaux. Les résultats, très prometteurs, seront détaillés plus loin. Nous présenterons d'abord le résultat de l'approche globale sur les lettres, puis son application en reconnaissance analytique de mots.

2 Rappel sur les champs de Markov causaux

Pour l'analyse d'une image, nous nous appuyons sur le NSHP-HMM introduit par George SAON ([3]). Ce modèle stochastique se base sur une décomposition selon les colonnes du calcul de la probabilité de l'image. La probabilité d'un point dépend d'un voisinage 2D pris dans le demi-plan non symétrique déjà évalué (Non-Symmetric Half Plan), liant les colonnes entre elles. L'élasticité horizontale du modèle est assurée par un HMM synthétisant l'analyse des colonnes dans ses états. La taille des colonnes est fixe pour un modèle, nécessitant une normalisation des échantillons analysés. Les paramètres du modèle sont le nombre d'états du HMM, la taille et forme du voisinage, la hauteur des colonnes analysées (hauteur du modèle). Nous utilisons un NSHP-HMM par classe d'images.

Par simplification, nous considérons un même voisinage (taille et forme) et une même hauteur pour toutes les classes d'images. Le nombre d'états du NSHP-HMM associé à une classe est obtenu en divisant par 2 la longueur moyenne des échantillons d'apprentissage de cette classe après normalisation proportionnelle.

3 Reconnaissance globale de lettres

Nous avons tout d'abord appliqué les NSHP-HMMs directement à la reconnaissance des lettres. Les premiers résultats sont de 50.45%. Constatant qu'une cause de confusion est une mauvaise répartition des états des NSHP-HMMs, nous avons introduit une transition de sortie de modèle dans les HMMs sous-jacents. Cette transition reflète la probabilité de sortir du modèle en chaque état. Cette amélioration nous a permis d'accroître le score de reconnaissance à 53.92%.

Le principal défaut du NSHP-HMM est sa rigidité verticale. Or la variabilité de l'écriture, le bruit et les caractéristiques des mots conduisent à des décalages verticaux importants, réduisant la correspondance

entre les images des lettres. Pour pallier ce problème, nous avons appliqué une normalisation par bandes des mots avant extraction des lettres. Cette normalisation consiste à rechercher les limites de la bande centrale par seuillage d'un histogramme combiné de transitions et de projection, puis à transférer les 3 bandes ainsi délimitées chacune sur 1/3 de l'image normalisée. Nous obtenons alors 60.59% de reconnaissance sur les lettres.

Les modèles utilisés pour ces résultats ont 40 pixels de haut et un voisinage de 3 points. Les tests ont portés sur 21183 échantillons de lettres segmentées de 4158 mots de montants de chèques provenant de la SRTP. Nous avons utilisé 90% de la base pour l'apprentissage et 10% pour la reconnaissance. La segmentation se fait par coupure en bandes verticales. Le vocabulaire est de 26 mots, composés de 21 lettres.

4 Reconnaissance analytique de mots

Nous avons étudié 2 approches analytiques, toutes deux guidées par le vocabulaire. La première se base sur une découpe en bandes verticales de l'image analysée. Chaque bande est associée à une lettre du mot supposé. Le NSHP-HMM associé à la lettre estime la vraisemblance de la bande, une gaussienne modélisant la probabilité de longueur de la bande sachant la lettre. Le système cherche par programmation dynamique les frontières les plus vraisemblables, qui donnent la vraisemblance du mot pour l'image.

La seconde approche consiste à lier les HMMs sous-tendant les NSHP-HMMs des modèles de lettres afin de générer des modèles de mots. Les modèles ainsi générés sont des NSHP-HMMs, donc s'appliquent directement, sans recherche des limites des lettres. Cependant, il n'est pas possible d'intégrer la notion de probabilité de durée des lettres dans ces modèles.

Pour ces deux approches, les NSHP-HMMs associés aux lettres ont un voisinage de 3 points et une hauteur de 20 pixels. La classification se fait par discrimination des résultats obtenus pour chaque mot du vocabulaire. Nous obtenons sur les mots 86.26% de reconnaissance pour la première méthode, et 84.22% pour la seconde.

Du point de vue de la complexité, l'approche analytique permet une réduction importante par rapport à l'approche globale proposée par George SAON ([3]). En effet, la complexité en calculs du NSHP-HMM est proportionnelle au nombre d'états du HMM sous-jacent: lors de l'estimation de la probabilité d'une image par un NSHP-HMM, chaque état du HMM sous-jacent analyse chaque colonne de l'image. Sur notre base, une approche globale des mots donne lieu à 26 modèles et un total de 615 états. L'approche analytique nous permet de réduire ces valeurs à 21 modèles totalisant 87 états. La complexité est donc réduite d'un facteur 7 entre les 2 approches.

5 Conclusion

Nous avons montré que l'approche analytique basée sur les NSHP-HMMs est performante, et permet de réduire la complexité de mise en oeuvre de ces modèles.

La validation de notre méthode demande des tests plus approfondis, sur des bases plus conséquentes.

Les modèles générés, bien que légèrement moins performants, présentent des cotés intéressants par leur simplicité de mise en oeuvre. Il serait intéressant de pouvoir récupérer la segmentation à partir de ces modèles, ce qui n'est pas possible actuellement.

Références

- [1] D. Guillevic and C. Suen. HMM Word Recognition Engine. In *Fourth International Conference on Document Analysis and Recognition (ICDAR'97)*, volume 2, pages 544–547, Ulm, Germany, Aug. 1997.
- [2] M. Gilloux, B. Lemarié, and M. Leroux. A Hybrid Radial Basis Function Network/Hidden Markov Model Handwritten Word Recognition System. In *Third International Conference on Document Analysis and Recognition (ICDAR'95)*, pages 394–397, Montréal, 1995.
- [3] G. Saon. *Modèles markoviens uni- et bidimensionnels pour la reconnaissance de l'écriture manuscrite hors-ligne*. PhD thesis, Université Henri Poincaré - Nancy I, Vandœuvre-lès-Nancy, 1997.