

Reconnaissance de l'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé

Laëtitia Rousseau, Eric Anquetil, Jean Camillerapp

► **To cite this version:**

Laëtitia Rousseau, Eric Anquetil, Jean Camillerapp. Reconnaissance de l'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé. MajecSTIC 2005: Manifestation des Jeunes Chercheurs francophones dans les domaines des STIC, IRISA – IETR – LTSI, Nov 2005, Rennes, France. pp.401-404. inria-00000740

HAL Id: inria-00000740

<https://hal.inria.fr/inria-00000740>

Submitted on 15 Nov 2005

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.

Reconnaissance de l'écriture manuscrite hors-ligne par reconstruction de l'ordre du tracé.

Laëtitia Rousseau, Éric Anquetil, Jean Camillerapp

Laboratoire IRISA - Campus Universitaire de Beaulieu

Avenue du Général Leclerc - 35042 RENNES Cedex - France

laetitia.rousseau@irisa.fr, eric.anquetil@irisa.fr, jean.camillerapp@irisa.fr

Résumé: La reconnaissance hors-ligne de l'écriture manuscrite consiste à interpréter ce qui a été écrit dans un document scanné. Cette technique est utilisée par exemple pour trier automatiquement le courrier. La reconnaissance en-ligne de l'écriture manuscrite consiste à interpréter une écriture représentée par la trajectoire du stylo. Cette technique est utilisée notamment dans les agendas électroniques. Les taux de reconnaissance sont meilleurs pour la reconnaissance en-ligne car les informations temporelles sont absentes des images. Retrouver l'ordre du tracé permettrait de traiter un document hors-ligne avec un reconnaisseur en-ligne et peut-être d'améliorer les résultats de la reconnaissance hors-ligne. Ce papier présente une approche originale de reconnaissance d'écriture hors-ligne, qui consiste à injecter des connaissances *a priori* sur l'écriture manuscrite pour reconstruire un signal en-ligne équivalent à l'image. Un certain nombre de connaissances sur l'écriture manuscrite utilisables pour retrouver l'ordre du tracé y sont analysées. Il s'agit de déterminer quelles connaissances peuvent être exploitées, comment les injecter et quelle est leur fiabilité. Ces premiers travaux ont été testés sur des lettres isolées et utilisés pour l'indexation automatique de documents.

Mots clés: reconnaissance de formes, écriture manuscrite hors-ligne et en-ligne, reconstruction de l'ordre du tracé.

1 INTRODUCTION

Depuis les années soixante, l'Homme cherche à "apprendre à lire" aux ordinateurs. Les caractères d'imprimerie ont tout d'abord été reconnus, puis les caractères manuscrits isolés, plus difficiles à reconnaître car leurs formes sont plus variées. Dans le cas des mots manuscrit cursifs, il existe des difficultés supplémentaires dues au fait que les lettres sont souvent attachées entre elles. Un problème majeur est de segmenter les mots en lettres. Pour les applications dont le vocabulaire est restreint et connu à l'avance, comme par exemple la lecture de montants de chèques, la reconnaissance est plus facile. Le mot à reconnaître est comparé aux mots de ce vocabulaire et la réponse du reconnaisseur correspond au mot le plus res-

semblant.

Il existe deux types de reconnaissance d'écriture. La reconnaissance hors-ligne permet d'interpréter ce qui a été écrit sur un document scanné. Cette technique est utilisée notamment pour trier le courrier, lire les montants de chèques... Le signal hors-ligne est une image en noir et blanc ou en niveaux de gris. Les méthodes généralement utilisées reposent sur des modèles de Markov cachés [Belaïd, 1997].

La reconnaissance en-ligne, quant à elle, permet d'interpréter une écriture représentée par la trajectoire du stylo. Cette technique est notamment utilisée dans les agendas électroniques de type PDA. Une tablette électronique et un stylo spécial sont nécessaires. Le signal est capté en temps réel. Il est constitué d'une suite de coordonnées de points, correspondant à la position du stylo à intervalles réguliers de temps.

La reconnaissance en-ligne possède de meilleurs taux de reconnaissance. En effet, le signal en-ligne contient des informations dynamiques absentes dans les signaux hors-ligne, telles que l'ordre dans lequel les caractères ont été formés, leur sens de parcours, la position des poser et lever de crayon... Retrouver l'ordre du tracé permettrait de traiter un document hors-ligne avec un reconnaisseur en-ligne dans l'espoir d'améliorer les taux de reconnaissance hors-ligne. Des connaissances *a priori* sur l'écriture manuscrite doivent alors être utilisées pour reconstruire un signal en-ligne équivalent à l'image. Il est important de souligner que l'objectif est de reconstituer un signal en-ligne cohérent qui n'est pas forcément celui qui a été réellement produit par le scripteur. En effet, des tracés différents peuvent produire une même image. Il suffit de retrouver un tracé interprétable par un reconnaisseur d'écriture.

La section suivante présente notre méthode de recherche de l'ordre du tracé. Plusieurs connaissances sur l'écriture sont analysées et des premiers résultats sur des lettres isolées sont présentés. Dans la partie trois, l'application à l'indexation automatique de documents est décrite.

2 RECHERCHE DE L'ORDRE DU TRACÉ

2.1 Principes de fonctionnement

Une image en niveaux de gris correspond à une matrice de luminosité de pixels selon les axes x et y . Binariser l'image signifie que l'on sépare ce qui fait partie du tracé et ce qui fait partie du fond. On cherche ensuite à squelettiser le tracé, c'est-à-dire à en extraire une ligne très fine correspondant à la trajectoire du stylo. Le problème est que le squelette est difficile à trouver au niveau des intersections. Pour y remédier, les lettres peuvent être représentées sous forme de zones régulières, représentant les portions longilignes donc squelettisables du tracé, et de zones singulières, correspondant aux régions difficilement squelettisables, telles que les intersections ou les fins de tracé (voir figure 1(b)). L'extracteur de tracé développé par l'équipe IMADOC¹ de l'IRISA, qui est une amélioration de [Pettier, 1993], permet d'obtenir cette représentation.

Une fois l'ordre du tracé reconstitué, le signal en-ligne correspondant (voir figure 1(c)) est généré et présenté à un reconnaiseur en-ligne (RESIFCar [Anquetil, 1997]). Pour déterminer la lettre lue, une comparaison avec les lettres de référence doit être effectuée. Pour cela, un apprentissage a préalablement été fait, durant lequel des signaux en-ligne ont été présentés au système en précisant de quelle lettre il s'agissait. Des caractéristiques ont alors été calculées pour chaque modèle de lettre. Quand une lettre inconnue est présentée au reconnaiseur, la réponse de celui-ci correspond au modèle qui lui paraît le plus proche.

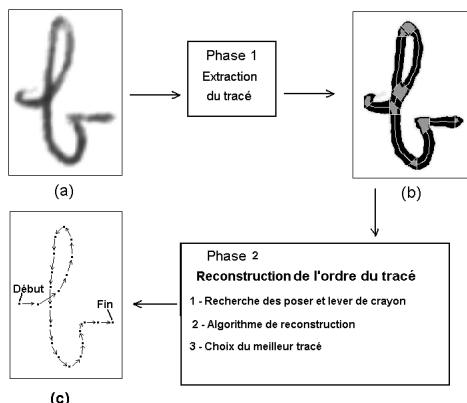


FIG. 1 – Représentations d'une lettre : (a) image en niveaux de gris (b) sous forme de zones régulières (en noir) et singulières (en gris) (c) suite ordonnée de points

2.2 Critères pour retrouver l'ordre du tracé

Les tracés qui doivent être reconstruits ne sont pas quelconques. Des connaissances *a priori* sur l'écriture manuscrite peuvent donc nous guider pour reconstituer leur ordre. Mais quels critères utiliser et comment les injecter ?

Dans [Boccignone, 1993], les critères utilisés sont locaux. Pour chaque intersection, les traits les plus alignés sont reliés entre eux. Dans [Jäger, 1996], les critères sont globaux. Tous les tracés possibles sont générés. Le tracé retenu correspond à celui dont la courbure globale et la longueur totale sont les plus faibles. Doermann et Rosenfeld [Doermann, 1995] utilisent des critères locaux et globaux. Ils introduisent notamment un critère lié au sens de parcours des boucles. En général, les boucles dirigées vers le bas sont tracées dans le sens horaire et celles dirigées vers le haut dans l'autre sens.

Notre approche consiste à injecter davantage de connaissances *a priori* sur l'écriture manuscrite. Les autres connaissances sur l'écriture utilisées sont les suivantes :

- Le tracé se fait de la gauche vers la droite (seule l'écriture occidentale est considérée ici).
- Une partie très courbe ne peut pas être tracée plusieurs fois.
- Les traits descendants sont les structures fondamentales de l'écriture et doivent être respectés.
- La direction la plus aisée pour écrire est la verticale descendante et la moins aisée est l'horizontale vers la gauche.

Nous avons évalué les critères ci-dessus. Pour cela, nous avons utilisé une base de lettres expérimentale [Viard-Gaudin, 1999].

2.3 Description de la méthode

Notre méthode pour retrouver l'ordre du tracé consiste d'abord à rechercher les poser et lever de crayon. Notre étude porte pour l'instant sur des tracés faits en une seule fois. Il n'existe donc qu'un poser et qu'un lever. A ce niveau de la chaîne, ils ne peuvent pas être déterminés de façon certaine, c'est pourquoi plusieurs candidats sont proposés. Pour chaque couple de candidats, le tracé correspondant est généré. Enfin, le meilleur tracé est conservé.

2.3.1 Recherche des poser et lever de crayon

La première étape consiste à localiser les poser et lever de crayon potentiels. Les deux extrémités de toutes les zones régulières sont des candidats potentiels. A partir de cet ensemble de poser ou lever potentiels, toutes les possibilités sont explorées.

Pour chaque couple de poser/lever potentiel, le sens de parcours est recherché. Pour cela, le fait que le tracé est en général effectué de la gauche vers la droite peut être utilisé. Ce critère mérite d'être raffiné. En effet, certaines lettres, comme le "s script", sont tracées de la droite vers la gauche. Mais dans ce cas, elles sont toujours tracées du haut vers le bas. Le principe décrit par la figure 2 permet de choisir le sens de parcours. La direction du vecteur reliant les deux points est calculée. Si cette direction est en bas à droite, A est le poser. Si c'est en haut à gauche, A est le lever. La partie hachurée correspond aux directions ne permettant pas de déduire le sens de parcours. Dans ce cas, les deux sens sont proposés.

1. site internet : www.irisa.fr/imadoc

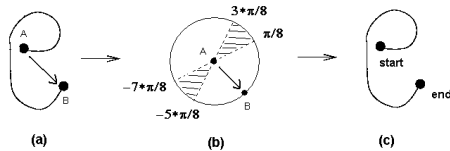


FIG. 2 – Détermination du sens de parcours

2.3.2 Algorithme de reconstruction

Une fois les poser et lever choisis, le tracé entre les deux doit être déterminé. A chaque intersection, il s'agit de relier les bons traits. Notre méthode s'est inspirée de l'algorithme de Kato et Yasuhara [Kato, 2000] dont nous avons proposé une extension [Rousseau, 2004]. Le principe de base est qu'il faut parcourir toutes les zones régulières. Dans un premier temps, les zones à parcourir plusieurs fois sont détectées, puis le tracé est recherché du poser au lever de crayon. Le choix au niveau des intersections se fait en choisissant le trait provoquant le plus faible changement de direction.

2.3.3 Choix du meilleur tracé

Une fois le tracé retrouvé pour chaque couple de poser et lever, il reste à déterminer quel est le meilleur de tous les tracés générés, en utilisant les critères décrits plus haut. Ces critères sont de deux types. Certains, comme le respect du sens de parcours des boucles, ne dépendent que du tracé considéré. Ces critères sont absolus. Les autres critères, comme la minimisation de la courbure globale, demandent de comparer les tracés entre eux. Ces critères sont relatifs.

Dans un premier temps, les tracés ne respectant pas les critères absolus sont éliminés. La courbure des parties parcourues plusieurs fois est calculée. Si elle dépasse un certain seuil, ce tracé est éliminé.

Le critère sur le sens des boucles n'est pas toujours fiable. La boucle dirigée vers le bas d'un f peut être faite dans le sens anti-horaire. Nous avons également étudié le sens de parcours des boucles dirigées vers la gauche ou vers la droite. Il s'avère qu'une boucle dirigée vers la gauche est parcourue dans le sens horaire et une boucle dirigée vers la droite est parcourue dans l'autre sens. Les tracés ne respectant pas le sens de parcours des boucles dirigées vers le haut, la droite ou la gauche sont donc éliminés.

Sur l'exemple de la figure 3, les deux tracés éliminés contenaient une partie très courbe parcourue deux fois.

Dans un second temps, les critères relatifs sont utilisés pour les tracés restants. La direction la plus aisée pour écrire est la verticale descendante et la moins aisée est l'horizontale vers la gauche. Un indicateur sur l'aisance d'écriture est calculé. La distance couverte dans une direction est multipliée par un coefficient dépendant de cette direction. La valeur maximum correspond à la direction "vers le bas" et la plus faible valeur à la direction "vers la gauche". Entre ces deux directions, la valeur croît de façon continue. Il s'agit de maximiser ce critère que nous appellerons "aisance".

Pour chaque tracé, la hauteur du plus grand trait descen-

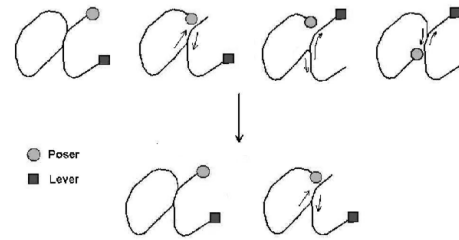


FIG. 3 – Choix du meilleur tracé. En haut : tracés générés. En bas : tracés conservés.

nant est calculée. Le tracé possédant la plus grande valeur a de grandes chances d'être le meilleur, car cela signifie que la zone descendante principale n'a pas été cassée. Il s'agit de maximiser ce critère que nous appellerons "descendants". Les autres critères testés sont la minimisation de la courbure globale et de la longueur totale.

Le tableau 1 rapporte le taux de fiabilité pour chaque critère utilisé seul. Ces taux ont été calculés en comptant le nombre de fois où le bon tracé satisfaisait le plus le critère. Par exemple, le bon tracé est le moins courbe dans 82,2% des cas. Les cas d'égalité ne sont pas comptés. En effet, le but est de proposer un seul tracé au reconaisseur en-ligne. Aucun critère ci-dessus n'est fiable à

courbure	aisance	descendants	longueur
82,2%	78,2%	73,2%	71%

TAB. 1 – Fiabilité des critères.

100%. C'est pourquoi plusieurs critères doivent être utilisés pour en déduire le bon tracé. Nous avons décidé de conserver les critères possédant les trois meilleurs taux. Le tracé le moins courbe, le plus aisé à dessiner et ne cassant pas un trait descendant important est recherché. La principale difficulté réside dans la façon de fusionner ces critères. Pour effectuer cette fusion, un classifieur de type réseau de neurones a été utilisé [Bishop, 1995]. Son principe est le suivant : lors de la phase d'apprentissage, les valeurs des trois critères sont présentées en précisant s'il s'agit d'un bon ou d'un mauvais tracé. Lors de la phase de test, seules les valeurs des critères sont présentées au classifieur qui répond par un coefficient d'appartenance à un bon tracé. Le tracé possédant le meilleur score est alors conservé. Le classifieur choisit le bon tracé dans 90% des cas.

2.4 Validation sur des lettres isolées

Nos tests portent sur la base de lettres minuscules IRO-NOFF [Viard-Gaudin, 1999] dont on possède également le signal en-ligne correspondant, ce qui permet de comparer les résultats des deux types de reconnaissance.

Les poser et lever de crayon sont parfois cachés dans le tracé, comme montré sur la figure 4. Pour cette première évaluation, nous avons enlevé ces lettres. Nous avons testé notre méthode sur des lettres tracées en une seule fois, dont les modèles ont été appris par notre reconnais-

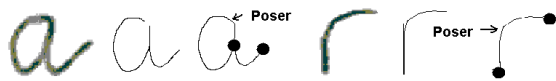


FIG. 4 – Poser et lever de crayon cachés

seur en-ligne et dont les poser et lever ne sont pas cachés, soit 5 800 lettres (sur les 10 000 contenues dans la base). Le signal en-ligne reconstitué est reconnu dans 80% des lettres, contre 87% pour le véritable signal en-ligne.

3 APPLICATION À L'INDEXATION DE DOCUMENTS

Les généalogistes sont en général amenés à consulter des milliers de documents. Certains ont été récemment numérisés, mais le changement de support ne modifie pas le nombre de consultations nécessaires. L'idéal serait d'avoir un logiciel qui, à partir d'un champ manuscrit saisi par un utilisateur, proposerait le document contenant ce champ ou du moins un petit nombre de documents. Pour cela, nous pouvons chercher à reconnaître l'écriture et associer la liste de lettres reconnues, appelée index, à chaque fichier.

Les documents que nous traitons sont des formulaires contenant un nom de famille à indexer. L'écriture, appelée écriture bâtarde, contient des lettres bien formées (voir figure 5). Après localisation dans le document du



FIG. 5 – Écriture bâtarde

nom de famille, nous recherchons l'ordre du tracé. Les pleins et les déliés étant très accentués dans ce style d'écriture, nous arrivons relativement bien à segmenter les mots en lettres sur un critère de différence de largeur du tracé. Puis nous appliquons la méthode décrite plus haut. Un signal en-ligne équivalent à l'image est alors généré et présenté au reconnaiseur en-ligne, qui propose en retour une suite de lettres. Cette suite de lettres compose l'index du document. Lorsqu'un utilisateur cherche un document en saisissant un nom de famille, tous les index sont comparés à sa requête et les images sont affichées par ordre de ressemblance.

Nos premiers tests sur trois cent trente documents sont encourageants. Le bon document est proposé en première position dans 50% des cas, et dans les dix premières positions dans 80% des cas.

4 CONCLUSION

Nous venons de présenter une approche originale de reconnaissance d'écriture hors-ligne, consistant à reconstituer l'ordre du tracé et à utiliser un reconnaiseur en-ligne. Un certain nombre de connaissances *a priori* sur l'écriture manuscrite ont été analysées et utilisées pour retrouver l'ordre du tracé. Seules les lettres isolées faites en une seule fois et les mots dont la segmentation est relativement facile ont été testés pour l'instant. La méthode va être étendue pour traiter des mots dont la largeur du tracé est relativement constante. D'autres connaissances sur l'écriture concernant plus particulièrement les mots seront étudiées.

BIBLIOGRAPHIE

- [Anquetil, 1997] Anquetil E. et Lorette G., On-line handwriting character recognition system based on hierarchical qualitative fuzzy modeling. *Progress in Handwriting Recognition*, pages 109–116. World Scientific.
- [Belaïd, 1997] Belaïd A. et Saon G., Utilisation des processus markoviens en reconnaissance de l'écriture. *Revue Traitement du Signal*, 14(2):161–177.
- [Bishop, 1995] Bishop C., *Neural Network for Pattern Recognition*. Oxford University Press.
- [Boccignone, 1993] Boccignone G., Chianese A., Cordella L. P. et Marcelli A., Recovering dynamic information from static handwriting. *Pattern Recognition*, 26(3):409–418.
- [Doermann, 1995] Doermann D. S. et Rosenfeld A., Recovery of temporal information from static images handwriting. *International Journal of Computer Vision*, 15:142–164.
- [Jäger, 1996] Jäger S., Recovering writing trace in offline handwriting recognition: Using a global optimization technique. *Proceedings of ICPR '96*, 3:150.
- [Kato, 2000] Kato Y. et Yasuhara M., Recovery of drawing order from single-stroke handwriting images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:938–949.
- [Pettier, 1993] Pettier J. et Camillerapp J., Script representation by a generalized skeleton. *Proceedings of the 2nd International Conference on Document Analysis and Recognition, (ICDAR'93)*, pages 850–853, Tsukuba, Japan.
- [Rousseau, 2004] Rousseau L., Anquetil E. et Camillerapp J., Reconstitution du parcours du tracé manuscrit hors-ligne de caractères isolés. *8ème Colloque International Francophone sur l'écrit et le Document, (CIFED'04)*, pages 123–127, La Rochelle, France.
- [Viard-Gaudin, 1999] Viard-Gaudin C., Lallican P. M., Knerr S. et Binter P., The ireste on/off (ironoff) dual handwriting database. *Proceedings of the Fifth International Conference on Document Analysis and Recognition (ICDAR'99)*, pages 455–458.