

Reconstruction de la Chronologie d'un Tracé Manuscrit Hors-Ligne

Pierre-Michel LALLICAN, Christian VIARD-GAUDIN

Equipe IVC
Laboratoire SEI EP0063 CNRS
IRESTE
Rue Christian Pauc
BP 60601
44306 NANTES Cédex 03

RÉSUMÉ

La reconnaissance de l'écriture manuscrite non contrainte reste confrontée à des problèmes très difficiles. Cependant, les méthodes de reconnaissance «en-ligne» présentent des résultats meilleurs que les approches «hors-ligne» qui perdent toute information temporelle. L'objectif du travail présenté ici est de restaurer l'ordonnement des éléments du tracé à partir d'images statiques 2-D, information qui est directement disponible dans les systèmes en-ligne. L'approche présentée ici est innovante, premièrement, contrairement à la plupart des autres méthodes, elle utilise au maximum l'information niveaux de gris, deuxièmement, elle utilise une approche par filtrage dans la prédiction de la trajectoire d'un élément du tracé.

1 Introduction

Les systèmes de reconnaissance de caractères manuscrits sont typiquement divisés en deux domaines bien séparés : les systèmes de reconnaissance de caractères en-ligne ou dynamique et ceux de reconnaissance hors-ligne ou statique. Cette différenciation est évidente si l'on considère le mode d'acquisition utilisé dans les deux domaines. Pour la reconnaissance en-ligne, les données sont capturées à l'aide d'une tablette électronique et sont enregistrées durant le processus d'écriture. L'écriture en-ligne se caractérise par deux vecteurs mono-dimensionnels $x(t)$ et $y(t)$ formés des coordonnées échantillonnées de la trajectoire du stylo. L'écriture hors-ligne, quant à elle, se traduit par une bitmap produite soit à partir d'une caméra soit d'un scanner, qui digitalise un document initialement écrit. Le manque d'information temporelle dans l'approche hors-ligne augmente bien évidemment les difficultés pour segmenter correctement les mots en caractères. Par exemple, des caractères se chevauchant spatialement ne posent pas de problème de segmentation pour les approches en-ligne puisqu'ils sont séparés dans le temps. Ainsi, il a été montré [5] que la reconnaissance manuscrite en-ligne présente des résultats supérieurs à celle dite hors-ligne, cette supériorité est d'autant plus marquée que l'écriture est cursive.

Il est donc intéressant de transformer un problème hors-ligne en un problème pseudo en-ligne. En réalité, toutes les informations dynamiques ne pourront être restaurées à partir d'une image statique ; l'accélération et la vitesse seraient

ABSTRACT

Non constrained handwriting recognition is still faced with very difficult problems. However, on-line recognition methods exhibit better results than off-line methods which lose all temporal information. The aim of the work presented here is to recover the strokes ordering from static 2-D images as it is inherently available from on-line systems. The approach presented here is innovative, first conversely to most other ones, it uses extensively the gray-level information, secondly because of the use of a filtering approach in the prediction of the writing strokes trajectory.

difficiles, voire impossibles, à restituer. Un objectif plus réaliste est de restaurer les séquences ordonnées des primitives constituant les mots manuscrits. Les principales difficultés rencontrées dans ce problème concernent les boucles et les retracés qui sont très communs dans l'écriture cursive.

Dans ce papier, nous abordons ce problème et proposons une approche innovante. D'une part, nous avons choisi de ne pas binariser, à l'inverse de la plupart des systèmes de traitement de l'écrit, l'image originale et par ailleurs nous combinons des informations de trajectoire et des caractéristiques morphologiques sur les traits pour définir un coût d'appariement entre deux éléments de trait.

2 Approche proposée

Partant d'une image d'entrée en niveaux de gris nous extrayons dans un premier temps les éléments correspondant à des traits manuscrits [4]. Une technique d'extraction des contours a été utilisée, elle est implémentée avec le filtre de Canny-Deriche. Une fois que tous les contours sont détectés dans l'image, chaque pixel est apparié avec le meilleur pixel correspondant sur le contour opposé. Quand aucun appariement n'est possible, le pixel appartient à une région non régulière (région singulière). Il résulte de cette procédure un ensemble de listes ordonnées de coupes transversales, une liste particulière correspondant à un sous-trait (segment) du tracé entre deux régions singulières (intersection ou extrémité de trait). Ce pré-traitement est décrit dans la section 3.

Un graphe composé de segments et de régions singulières est alors disponible. L'objectif est alors d'extraire de ce graphe les chemins optimaux représentant les parties du tracé entre un posé et un levé de stylo [3]. Cela consiste à étiqueter successivement toutes les régions singulières selon leur nature, ceci étant réalisé en étudiant la structure géométrique et morphologique des différents segments liés à chacune des régions dans le graphe. Une mesure de mise en correspondance entre trajectoires de segment est couplée à des caractéristiques sur les segments relevant notamment de leur largeur et de leur profil de luminosité, pour former un système décisionnel qui estime la nature de la région.

3 Segmentation des traits

La localisation des contours dans l'image en niveaux de gris est réalisée par l'extraction des passages par zéro du laplacien calculé par un filtre de Canny-Deriche implémenté de manière récursive [1]. Un seuillage avec hystérésis sur la norme du gradient permet de filtrer les passages par zéro du laplacien non significatifs. La figure 1 montre l'image en niveaux de gris (a) et les contours détectés correspondants (b). Ensuite, une coupe transversale germe est sélectionnée, elle doit satisfaire plusieurs critères : faible courbure des deux points contours extrémités, direction du gradient similaire pour ces deux points, largeur locale minimale de cette coupe transversale.

Une fois qu'une coupe germe est sélectionnée, l'extraction du segment est poursuivie avec une progression cohérente le long des deux bords, dans les deux directions. Cette progression est interrompue lorsqu'une région singulière est rencontrée, ceci se produit quand l'une des anomalies suivantes est rencontrée : une variation notable de la largeur, un profil de luminosité incorrect, une intersection avec un autre segment.

Dans la figure 1-b les frontières entre régions singulières et segments sont marquées avec une couleur de pixel plus foncée.

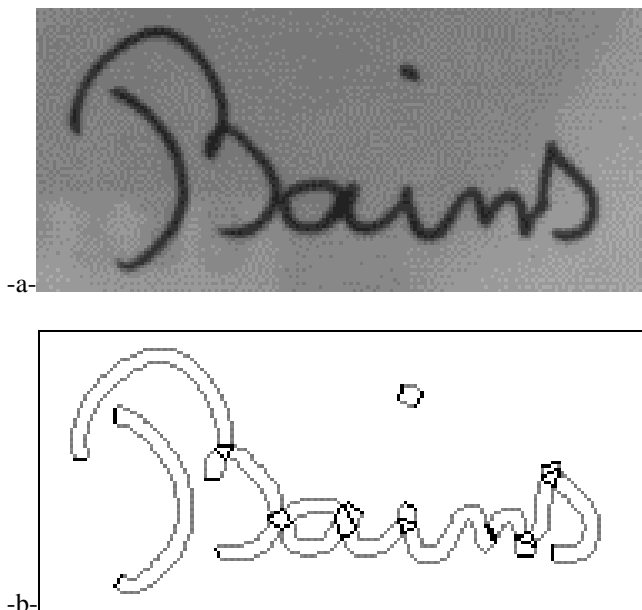


figure 1 : Extraction des contours et segmentation des traits

4 Interprétation et étiquetage des régions singulières

Au premier niveau de discrimination, le type d'intersection est classé en fonction du nombre de segment s'y joignant, les intersections les plus communes comportant 2, 3 ou 4 segments. Si 2 segments sont présents, différentes interprétations sont possibles, les plus courantes étant celle d'un point de forte courbure (fig.2.a) ou d'un coude (fig.2.b). D'autres interprétations peuvent être émises telles que la présence d'un retracé de stylo dans l'un des deux segments (fig.2.c), ou bien encore un contact accidentel entre deux extrémités de segment (fig.2.d). Si la région singulière comprend 3 segments, les deux principales hypothèses que nous retiendrons sont soit une fusion de deux segments en un troisième segment retracé (rebroussement - fig.2.e ou boucle fermée (fig.2.f), soit une intersection de type 'T' lorsqu'un segment se termine sur un autre trait (fig.2.g). Quant aux régions à 4 segments, un croisement de deux traits (fig.2.h) est le premier type d'intersection à considérer, des contacts entre deux traits sont également courants (fig.2.i). Des intersections plus complexes peuvent également être prises en compte apparaissant comme des combinaisons de jonctions fondamentales (fig.2.j et fig.2.k).

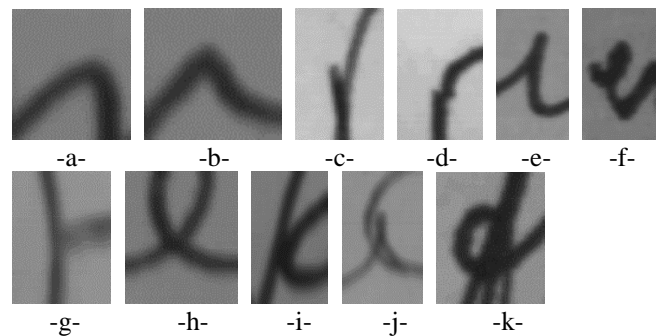


figure 2 : Régions singulières

4.1 Filtre d'estimation des paramètres d'une trajectoire

La principale hypothèse que nous faisons, est que la trajectoire peut être décrite par une succession d'arcs de cercle. Nous estimons itérativement la courbure de la trajectoire en chacun des points squelettes (points milieux des coupes transversales) en faisant l'hypothèse de continuité de position et d'orientation entre deux arcs de cercle consécutifs. La fenêtre d'estimation comprend un nombre de points squelettes situés dans un « futur » proche du point dont on estime la courbure. Cette fenêtre est de taille variable en fonction de la courbure estimée au point précédent, elle est grande dans une zone à courbure constante (meilleur filtrage) et diminue lorsque la courbure change.

4.2 Fonction de mise en correspondance entre deux trajectoires

La fonction de mise en correspondance est une mesure de la confiance qu'une paire de segments i et j sont portions d'un même trait étendu à travers une région singulière. Le principe pour calculer cette fonction, est basé sur les trajectoires de chacun des segments (position, orientation et courbure - voir Filtre d'estimation des paramètres d'une trajectoire §4.1). On prédit la trajectoire d'un des segments (segment i - voir figure 3) à travers la région singulière, on évalue alors la différence entre ce segment prédit et le segment restant (segment j). Le procédé est réitéré de la même façon en partant de l'autre segment (segment j vers segment i).

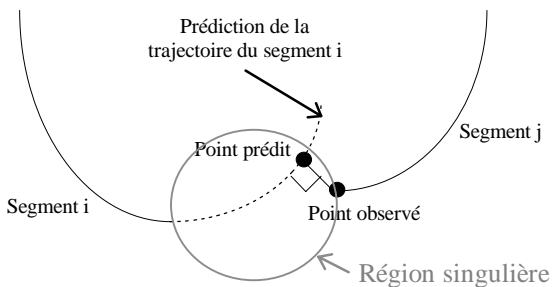


figure 3 : Prédiction de la trajectoire d'un segment

La différence entre les deux trajectoires se mesure au niveau du dernier point de la trajectoire du segment j (point observé) et de sa projection sur la trajectoire prédite du segment i . (point prédit). Elle prend en compte les erreurs de position, d'orientation et de courbure entre les deux points. Toutes ces erreurs peuvent être réunies dans une unique fonction de mise en correspondance que nous avons définie ici comme une combinaison linéaire des trois erreurs (voir équation 1).

$$\text{segm } i \rightarrow \text{segm } j = \alpha \cdot \varepsilon_{\text{pos}} + \beta \cdot \varepsilon_{\text{dir}} + \gamma \cdot \varepsilon_{\text{courb}} \quad (1)$$

Si l'on est intéressé par une fonction non directionnelle, une moyenne pondérée entre i vers j et j vers i peut être calculée. Cette pondération est dépendante de la confiance que l'on a dans le filtrage de chacune des trajectoires des segments i et j . Plus l'extrémité du segment i est à courbure constante, plus sa prédiction sera prise en compte, c'est à dire plus μ_i sera grand (équation 2).

$$\text{segm } i \leftrightarrow \text{segm } j = \mu_i(\text{segm } i \rightarrow \text{segm } j) + \mu_j(\text{segm } j \rightarrow \text{segm } i) \quad (2)$$

Les résultats de la fonction de souplesse sur l'image de la figure 1-a sont fournis par les tableaux suivants.

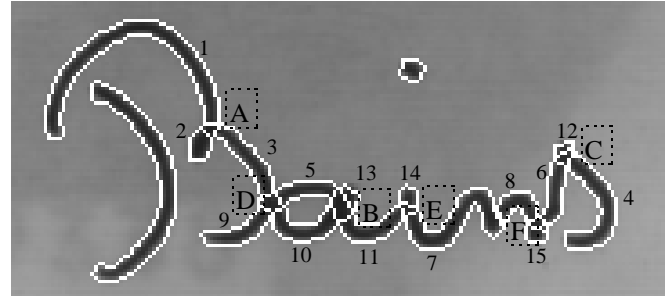


figure 4 : Numérotation (aléatoire) des segments

| région A | | | région B | | | | | |
|-------------|-------------|-------------|----------|-------------|-------------|-------------|-------------|-------------|
| 1-2 | 1-3 | 2-3 | 5-10 | 5-11 | 5-13 | 10-11 | 10-13 | 11-13 |
| 0.91 | 1.83 | 1.23 | 2.45 | 2.57 | 1.83 | 3.63 | 0.87 | 1.11 |
| région C | | | région D | | | | | |
| 4-6 | 4-12 | 6-12 | 3-5 | 3-9 | 3-10 | 5-9 | 5-10 | 9-10 |
| 2.81 | 1.09 | 0.42 | 2.78 | 0.17 | 1.38 | <u>1.27</u> | 1.32 | 4.08 |
| région E | | | région F | | | | | |
| 7-11 | 7-14 | 11-14 | 6-8 | 6-15 | 8-15 | | | |
| 3.38 | 0.66 | 0.58 | 3.25 | 1.17 | 0.87 | | | |

tableau 1 : Résultats de la fonction de mise en correspondance des segments de la figure 4

Ces tableaux nous montrent, par exemple, que si l'on considère la région D, la fonction de mise en correspondance est bien minimum pour l'association des deux paires de segments : 3 avec 9 et 5 avec 10 conduisant à valider l'hypothèse d'un contact entre deux traits vis à vis de celle d'un croisement ($3-9 + 5-10 < 3-10 + 5-9$). L'hypothèse de 3 et 10 en 9 (retracé du segment 9) reste à ce stade possible, mais sera invalidée par la suite lors de l'étude de la morphologie du segment 9.

4.3 Caractéristiques morphologiques des segments

Les mesures géométriques sur les trajectoires des traits servent de base à la construction d'un système décisionnel sur la nature des régions singulières. Cependant dans certains cas, ces mesures ne sont pas suffisantes pour prendre une décision correcte, des informations extraites des niveaux de gris de l'image et relatives à la morphologie des segments peuvent contribuer à un meilleur choix [2]. Par exemple, la largeur et le profil de luminance de deux éléments du tracé écrits consécutivement doivent être similaires. De même, un segment retracé présente une largeur et un contraste élevé dans le cas d'un rebroussement, et dans le cas d'une boucle fermée, son profil de luminance se compose de deux vallées (figure 5-b).

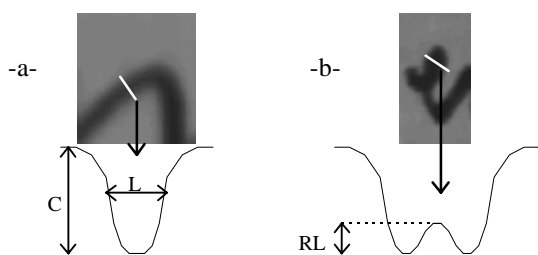


figure 5 : profil de luminance de coupes transversales

Nous définissons une mesure de retracé d'un segment par l'équation 3. Cette mesure est souhaitée faible lorsque le segment est effectivement un retracé et plus élevée sinon. Pour cela, un écart important de la largeur L maximale ou du contraste C maximal par rapport à leur valeur moyenne sur l'ensemble du tracé vont contribuer à une faible valeur de cette mesure. De la même façon, une remontée de luminance RL , ainsi qu'une faible courbure conduisent à une valeur faible pour la mesure de retracé. Cette dernière devra être également d'autant plus faible que la longueur du segment est grande. Ainsi, nous avons introduit dans l'équation 3 non pas une mesure de la courbure moyenne du segment, mais de la somme des courbures des points squelettes P_{sq} qui ont été initialement sous échantillonnés par rapport à la largeur moyenne du tracé.

$$ret(seg m) = \alpha \frac{L_{moy} - L_{max}}{L_{moy}} + \beta \frac{C_{moy} - C_{max}}{C_{moy}} + \gamma \frac{0,25 \cdot C_{moy} - RL_{max}}{0,25 \cdot C_{moy}} + \delta \sum \frac{^{seg m} courb(P_{sq})}{courb_{moy}} + offset \quad (3)$$

Finalement, une évaluation d'un posé ou d'un levé de stylo peut être calculée en chaque extrémité de segment (équation 4). Celle-ci fait l'hypothèse que le segment auquel appartient l'extrémité ne soit pas un retracé, et qu'au niveau de cette extrémité, si celle-ci est reliée à une région singulière comprenant au moins deux segments, le segment présente des fonctions de souplesse élevées avec les autres segments joints à cette région.

$$lev - pos(extrem i) = \frac{1}{ret(seg m i)} + \frac{1}{\min_j(seg m i \rightarrow seg m j)} \quad (4)$$

La figure 6 donne un aperçu des résultats que l'on obtient pour la mesure d'un retracé et celle d'un levé ou posé de stylo sur quelques segments de l'image étudiée. Ces mesures sont effectivement faibles dans les cas souhaités.

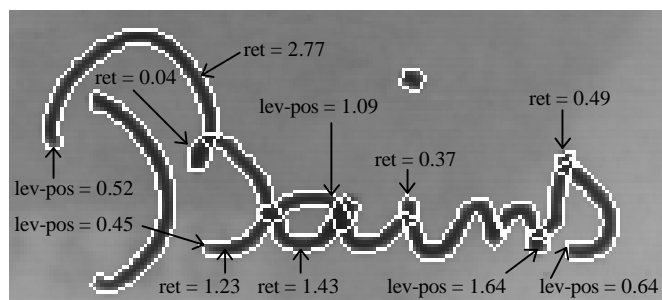


figure 6 : Mesures de retracé et de levé ou de posé de stylo

5 Conclusion

Les travaux présentés ici ont pour objectif final de définir une plate-forme qui caractérise complètement le tracé dans le contexte de la reconnaissance de l'écriture manuscrite, le premier but étant de restaurer les informations temporelles à partir d'une image statique. Pour cela, nous avons déjà obtenu des mesures locales robustes pour caractériser les segments et les régions singulières. Ceux-ci vont être intégrés dans un système plus global assurant la cohérence régionale des associations en prenant en particulier en compte les règles de formation de l'écriture.

Références

- [1] R. Deriche, «Optimal edge detection using recursive filtering», First International Conference on Computer Vision, Juin 1987, pp. 501-505.
- [2] D.S. Doermann, A. Rosenfeld, «The Interpretation and Reconstruction of Interfering Strokes», IWFHR'93, pp. 41-50.
- [3] S. Jäger, «Recovering writing traces in off-line handwriting recognition: using a global optimization technique», ICPR'96, pp. 150-154.
- [4] J.C. Pettier, «Extraction d'une représentation adaptée à la reconnaissance de l'écriture», Thèse, Université de Rennes I, IFSIC, 1994.
- [5] R. Seiler, M. Schenkel, F. Eggimann, «Off-line cursive handwriting recognition compared with on-line recognition», ICPR'96, pp. 505-509.