Une méthode de résolution des conflits d'alignements pour la segmentation des documents manuscrits

Solving Alignment Conflicts in Handwritten Document Segmentation

par Laurence LIKFORMAN–SULEM, Claudie FAURE
ENST URA-CNRS 820 Traitement et Communication de l'Information
46, rue Barrault, F-75634 Paris cedex 13

Résumé

La segmentation en lignes d'un document est une étape nécessaire avant d'aborder la reconnaissance des caractères, symboles ou mots. Cet article présente une méthode itérative basée sur le groupement perceptif, adaptée aux documents manuscrits non contraints. A partir de points d'ancrage directionnels, les composantes satisfaisant aux critères de proximité, similarité et continuité de direction sont groupées pour former des alignements. Les conflits qui apparaissent du fait de l'imbrication des alignements ou du chevauchement des hampes et jambages sont résolus soit localement par application du critère de continuité de direction, soit globalement en examinant la configuration et la qualité des alignements.

Mots clés : segmentation en lignes, document manuscrit, structuration perceptive.

Abstract

Text line segmentation is necessary before performing character recognition. We present here an iterative method for extracting text lines in unconstrained handwritten documents which uses clues from laws of perceptual organisation. Alignments are build from anchor points, linking components under criteria such as proximity, similarity and direction continuity. Conflicts may appear due to overlapping or interwoven lines. A local procedure guided by a direction continuity criteria, first seeks to solve the conflict. It may be followed by a global procedure which is based on the configuration of the alignments and their perceptual quality.

Key words : *line segmentation, handwritten document, perceptual organisation.*

1. Introduction

L'analyse des documents manuscrits se propose de lire et de convertir les pages de textes préalablement numérisées en une version ASCII manipulable par un traitement de texte. Cette lecture automatique, qui reste un problème ouvert, permettrait de réaliser des tâches fastidieuses telles que trier le courrier, créditer un compte bancaire, mettre au propre un brouillon, et d'offrir de nouveaux outils pour l'interrogation de bases de données notamment en recherchant un document à partir du contenu et non à partir de mots clés fixés par un opérateur humain.

Avant la reconnaissance des symboles ou des mots, les pages de texte doivent être segmentées en composantes physiques telles que les lignes, les mots ou parties de mots. Nous nous attachons dans cet article à la structuration en lignes des textes, dont la perception est indépendante du contenu linguistique.

Les stratégies d'extraction de lignes pour l'analyse des documents imprimés se divisent en deux catégories : descendantes et ascendantes. Les méthodes descendantes divisent le document pour aboutir aux lignes : ce sont les méthodes basées sur les projections [1]. Les méthodes ascendantes partent des pixels ou des composantes connexes de l'image et les fusionnent pour former des lignes. On peut citer le lissage en séquence de plage [2], la transformation de Hough [3], le regroupement de composantes connexes par proximité [4]. Une approche mixte [5] consiste à localiser les lignes à partir d'une représentation multi–résolution du texte. Ces méthodes ont été appliquées à des textes manuscrits dont la mise en page n'est pas trop éloignée de celle du document imprimé : lignes proches de l'horizontale, bien espacées ([6] [7] [8] [9] [10]).

L'écriture manuscrite présente de grandes variations, dans les formes des lettres ou des mots, et dans la mise en page. Les lignes sont de longueurs différentes. Les espacements entre lignes, entre mots, entre parties de mots sont irréguliers. Différentes directions de ligne peuvent coexister sur une même page du fait de changements d'orientation de la feuille en cours d'écriture, ou de l'insertion de notes en marge du texte. D'autre part les lignes de texte peuvent être imbriquées et même collées quand hampes et jambages appartenant à deux lignes consécutives sont proches ou

se chevauchent. Dans les textes manuscrits non contraints (brouillons, notes, enveloppes postales) cette situation est courante et met en défaut les procédures classiques de détection de lignes telles que les projections et le lissage en séquence de plage. Il devient donc nécessaire de concevoir des algorithmes de segmentation spécifiques au document manuscrit.

Nous avons présenté dans [11] une méthode itérative de segmentation en lignes des textes manuscrits à partir des composantes connexes de l'image. Cette méthode aspire à traiter les documents manuscrits dans leur généralité, sans hypothèses sur la direction *a priori* des lignes.
Le processus de segmentation que nous avons développé s'appuie sur les mécanismes perceptifs qui permettent à l'être humain de voir des lignes de texte, notamment à distance, indépendamment de la lecture proprement dite. Elle utilise certains principes de la physiologie de la vision et des lois d'organisation perceptive de la théorie de la Forme (Gestalt). A partir d'un ensemble d'excitations élémentaires produites par des unités d'écriture orientées dans une direction, les composantes de l'image sont organisées suivant des structures linéaires pouvant présenter de légères courbures ou des fluctuations autour d'une direction principale.

Le but de la segmentation est d'affecter à chaque composante connexe une étiquette unique, le numéro identificateur de la ligne à laquelle elle appartient. Pendant le processus de groupement, les situations de conflits apparaissent lorsqu'une composante appartient à plusieurs alignements. Ces situations apparaissent du fait de l'interpénétration des lignes ou de leur chevauchement. Ces composantes sont alors marquées comme ambiguës. Il s'agit d'affecter la composante à l'alignement qui réalise la meilleure forme, au sens de l'organisation perceptive. La meilleure forme est celle qui donnera avec la composante concernée, l'alignement de meilleure continuité lui assurant une plus forte pregnance perceptive.

Cet article rappelle les principes de groupement des composantes connexes en alignements, montre comment repérer les composantes ambiguës et réaliser une analyse locale du conflit qui, si elle est suffisante, permet d'affecter ces composantes à un alignement unique. Dans le cas contraire, l'analyse locale est suivie d'une analyse globale qui affine la détection des alignements.

2. Description générale

Le principe de la méthode de groupement des composantes en alignements est le suivant. On cherche tout d'abord à évaluer localement la ou les directions présentes dans la page : c'est le rôle de la détection des points d'ancrage. L'espace de la page est balayé à partir des points d'ancrage, suivant leur direction, à la recherche des composantes connexes formant des alignements. La méthode se décompose en plusieurs étapes :

1) recherche des composantes connexes
2) sélection des composantes connexes formant les points d'ancrage directionnels
3) construction des alignements par regroupement des composantes voisines
4) évaluation de la qualité des alignements
5) résolution des conflits
6) incrémentation du seuil de voisinage, retour en 3.

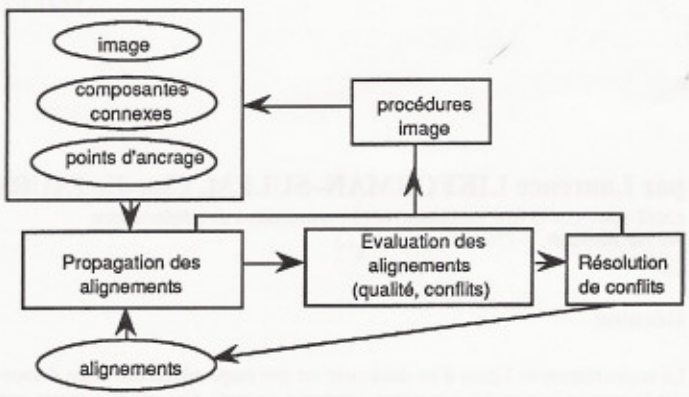

Fig. 1. – Stratégie de construction des alignements.

La première étape est la procédure classique d'étiquetage des composantes connexes. Celles ci sont des mots ou parties de mots, des caractères ou parties de caractères, ou un groupe de mots. La deuxième étape consiste à appliquer des masques sur les composantes pour sélectionner celles qui ont une direction fiable. A partir des points d'ancrage, on forme des alignements en regroupant les composantes connexes voisines dans un voisinage dépendant d'un seuil de proximité S. Les alignements obtenus sont évalués du point de vue de leur qualité et des conflits possibles. En cas de conflit, des règles sont appliquées pour les résoudre. La figure 1 illustre la boucle itérative et les différents modules qui la composent : les modules liés aux étapes 3, 4 et 5 construisent les alignements et forment une boucle dans laquelle le seuil de proximité S est incrémenté à chaque itération. Des procédures image sont attachées aux modules pour la détermination des relations de voisinage entre composantes.
La segmentation s'arrête quand les alignements ne sont pas modifiés entre deux itérations. Un post–traitement consiste, pour chaque composante restée isolée, à rechercher la composante la plus proche dans l'image et à lui affecter l'étiquette d'alignement correspondante, si elle existe.

3. Définitions

• *rectangle englobant*

Soit i une composante connexe, on note S_i son rectangle englobant

• *fonction de présence*

Soit la fonction $P(i, S)$ qui calcule le nombre de pixels de la composante connexe i dans la zone rectangulaire S de l'image.

Le nombre de pixels d'une composante i est : $P(i, S_i)$

• *hauteur d'un alignement*

Soit un alignement A. On note $H(A)$, la distance comprise entre les positions extrêmes (supérieures et inférieures) de l'alignement.

• *relations de voisinage*

Soit une composante connexe v. Lors du groupement, la composante c est trouvée voisine. Suivant le sens de balayage, la relation de voisinage s'exprime par la relation :

$$c = \text{voisin}_1^+(v) \quad \text{ou} \quad c = \text{voisin}_1^-(v)$$

Le chiffre 1 indique que le voisin trouvé est le *premier* plus proche voisin. Le signe $+$ ou $-$ indique le sens de balayage de l'espace de la page. Quand la direction de l'alignement est l'horizontale, le voisin$^+$ est le voisin de droite et le voisin$^-$, le voisin de gauche. Une fois le voisin trouvé, on établit la relation réciproque :

$$v = \text{voisin}_1^-(c) \quad \text{ou} \quad v = \text{voisin}_1^+(c)$$

Des voisins d'ordre deux peuvent être recherchés. Si d est la deuxième composante voisine de v, la relation s'exprime par $d = \text{voisin}_2^+(v)$.

4. Détection des points d'ancrage

Nous nous sommes inspirés du système visuel humain, notamment des propriétés des cellules simples de l'aire 17 du cortex visuel [12]. La configuration des champs récepteurs de certaines cellules les rendent sensibles à l'orientation des segments de lignes. Par analogie avec les champs récepteurs, nous définissons 4 masques correspondants à une discrétisation de l'espace en 4 directions (0°, 45°, 90°, 135°). Chaque masque m est carré et contient une zone d'excitation z_m^+ sensible à la densité d'écriture dans la zone [figure 2].

Ces masques sont appliqués sur toutes les composantes connexes, leur taille s'adaptant à celle de la composante. La réponse d'une composante i au masque m est :

$$R_m(i) = \frac{P(i, z_m^+)}{P(i, S_i)}$$

Les points d'ancrage sont des composantes connexes dont l'orientation est fiable (typiquement $R_m(i) \geq 80\%$). Ce sont les composantes qui ont une forme compacte et « allongée » le long d'une direction : segment d'écriture sans hampe ni jambage, lettre longiligne. L'orientation locale de ces composantes peut coïncider ou non avec la direction globale de l'alignement auquel elle appartient. La méthode est sensible à cette caractéristique et il faut

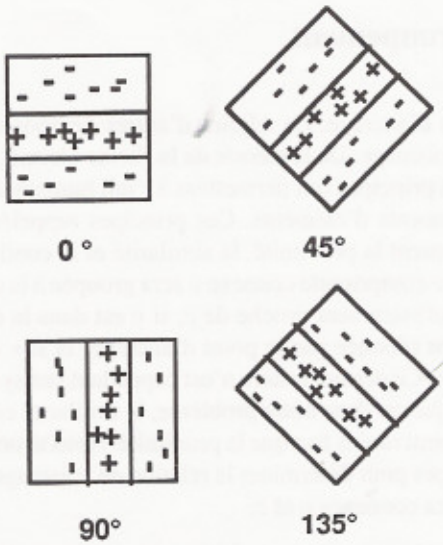

Fig. 2. – Masques de détection d'orientation suivant les 4 directions. La zone d'excitation z_m^+ est signalée par les signes $+$.

trouver suffisamment de points d'ancrage dont les orientations coïncident avec celles des lignes de texte dans lesquelles elles sont incluses.

Les composantes points d'ancrage sont les points de départ des alignements [figure 3].

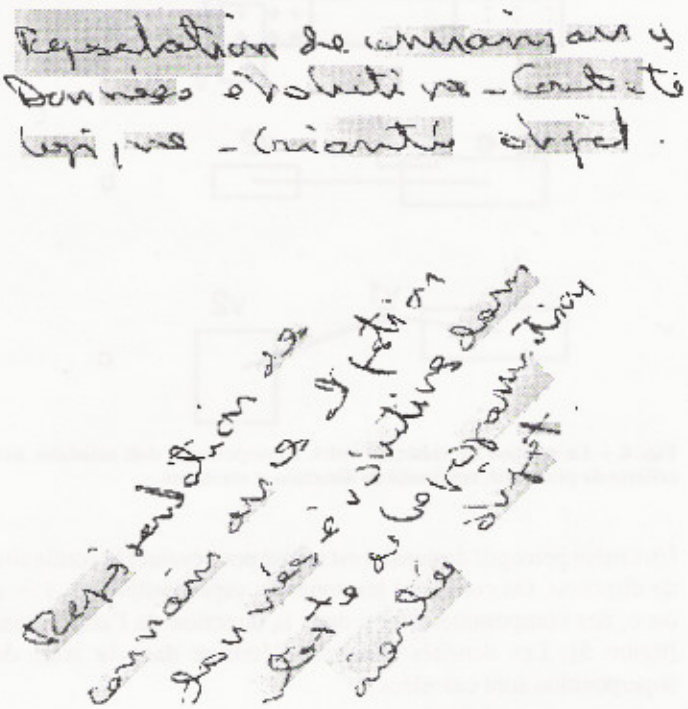

Fig. 3. – Les alignements sont construits à partir des points d'ancrage sélectionnés (en grisé).

5. Groupement

Aux points d'ancrage, on adjoint d'autres composantes situées dans leur voisinage. De la théorie de la Forme Gestalt, nous avons retenu trois principes qui permettent à l'œil humain de percevoir des groupements d'éléments. Ces principes rappelés dans [13] sont notamment la proximité, la similarité et la continuité de direction. Une composante connexe v sera groupée à la composante c si v est suffisamment proche de c, si v est dans la direction de l'alignement (donnée par le point d'ancrage) et si v est de taille similaire à c. Ce dernier critère n'est cependant pas systématiquement appliqué car dans notre problème, la similarité est un facteur de groupement moins fort que la proximité. Nous avons utilisé ces trois principes pour déterminer la relation de voisinage entre deux composantes connexes v et c.

5.1. PROXIMITÉ ET CONTINUITÉ DE DIRECTION

Pour satisfaire au critère de proximité v doit être à une distance $D \leq S(k)$, $S(k)$ étant le seuil maximal de groupement autorisé à l'itération k du processus [figure 4–a]. Suivant la position de v par rapport à c et la répartition de l'écriture à l'intérieur de chaque composante, v est accepté ou non comme voisin.

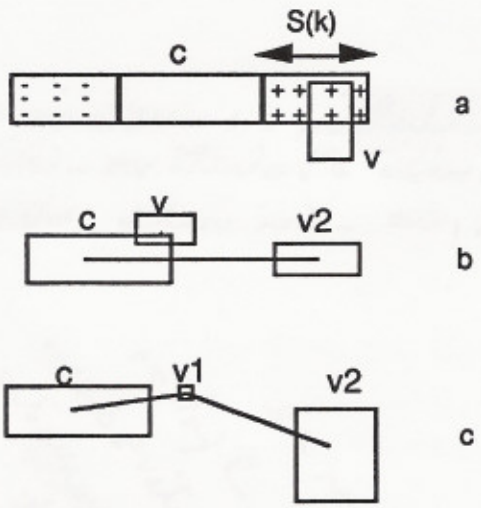

Fig. 4. – La relation de voisinage entre 2 composantes doit satisfaire aux critères de proximité, continuité de direction et similarité.

Un critère perceptif de densité est utilisé pour évaluer la continuité de direction. On considère les zones de superposition z_S^i, $i = c$ ou v, des composantes c et v dans la direction de l'alignement [figure 5]. Les densités relatives d'écriture dans la zone de superposition sont calculées :

$$R_i = \frac{P(i, z_S^i)}{P(i, S_i)} \quad i \in \{v, c\}$$

Intuitivement, si ce rapport a une valeur élevée, c'est que le corps de l'écriture de la composante i (qui correspond à la largeur de la bande centrale de l'écriture) est dans la zone de superposition. Si les deux rapports R_v et $R_{\not v}$ sont faibles tous les deux, v n'est pas accepté comme voisin de c [figure 4–b].

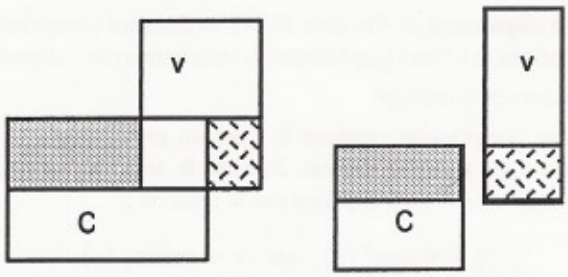

Fig. 5. – Zones de superposition z_s^c (grisé) et z_s^v (chevrons) suivant les positions des composantes c et v.

5.2. SIMILARITÉ

Le critère de similarité n'est utilisé que si l'on ne trouve pas de voisin à une composante trop petite. Il est estimé par le rapport des hauteurs des composantes connexes c et v. Si ce rapport est faible, les composantes sont considérées comme non similaires. Sur la figure 4–c, $v1$ est voisin de c, mais $v1$ n'a pas de voisin ($v1$ peut être un accent). Du fait de la non similarité de $v1$ et de c, on recherche alors un deuxième voisin à c, satisfaisant néanmoins aux critères de proximité et de densité. Le deuxième voisin, $v2$, s'il existe, sera alors considéré comme voisin de $v1$ dans l'alignement. Cette procédure permet de poursuivre les alignements, malgré la présence de petites composantes qui pourraient arrêter leur avance.

La relation de voisinage entre deux composantes v et c est notée par les deux relations :

$$v = \text{voisin}_1^+(c) \quad \text{et} \quad c = \text{voisin}_1^-(v)$$

Car quand on obtient une relation de voisinage, par exemple $v = \text{voisin}_1^+(c)$, on établit la relation réciproque.

Pour faciliter la compréhension de la suite de l'article, nous supposerons la direction horizontale. Le signe + indiquera le sens de balayage de l'espace de la page vers la droite, le signe −, le sens vers la gauche [figure 4–a].

Des voisins d'ordre supérieur peuvent aussi être recherchés en cas de conflits (cf. section 6).

5.3. FORMATION DES ALIGNEMENTS

Au départ, les alignements sont réduits aux points d'ancrage. A la première itération, l'image est balayée à partir de chaque composante connexe ayant été identifiée comme point d'ancrage, dans la direction donnée par l'orientation de cette composante. On

regroupe ainsi les composantes connexes voisines par propagation de l'information de direction. Les ensembles de composantes ainsi constitués forment des alignements encore partiels. A chaque nouvelle itération, le seuil de proximité S est augmenté et de nouvelles composantes sont ainsi adjointes aux alignements déjà constitués. Quand les alignements se stabilisent, la segmentation finale est obtenue.

6. Analyse locale des conflits

6.1. SITUATION DE CONFLITS

Lors de la formation des alignements, une composante déjà affectée à un groupement peut être rattachée au groupement courant : la composante est alors ambiguë. Un conflit apparaît entre les alignements concernés, ceux–ci pouvant être de directions différentes s'ils se croisent ou de même direction si l'un est en dessous de l'autre. Ces différentes situations sont notées : croisement, embranchement de type Y et embranchement de type U.

Les croisements [figure 6–a] sont dus aux lettres, symboles longilignes ou accents d'où partent des alignements transversaux. L'ambiguïté pour la composante X sera levée lors de l'analyse globale par évaluation de la qualité des alignements en conflits (cf. section 6).

Dans un embranchement en Y, il existe des composantes de part et d'autre de la composante X [figure 6–b]. Un embranchement en U est caractérisé par l'absence de composantes sur un côté de la composante ambiguë X : on a alors $c2 = \phi$. Les analyses qui vont suivre traitent la situation où $c1$ et $c3$ sont à gauche de X (figure 6–b). Les traitements relatifs à la situation symétrique ($c1$ et $c3$ à droite de X, $c2$ à gauche de X) s'en déduisent aisément.

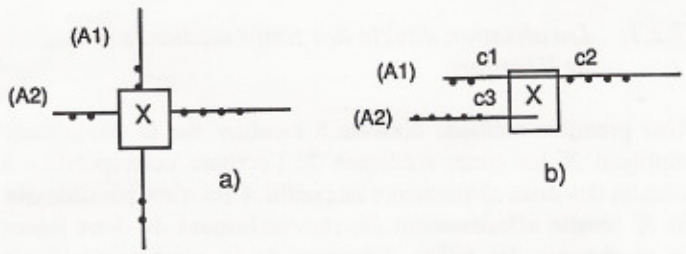

Fig. 6. – Situations de conflits. Les composantes dans les alignements sont représentées par des points, la composante ambiguë par un rectangle.

L'analyse locale des situations d'embranchements consiste à chercher le meilleur chemin pour les alignements concernés. Par application du principe de continuité de direction, on cherche à poursuivre les alignements au delà de la composante conflictuelle : des procédures de recherche de voisins sont appliquées à l'image pour trouver des voisins d'ordre supérieurs. L'analyse locale permet soit de régler directement le conflit, soit d'intégrer la composante aux deux alignements en conflit en signalant son ambiguïté.

6.2. ANALYSE DES EMBRANCHEMENTS

Soit la situation d'embranchement telle qu'elle est représentée figure 6–b, où deux alignements notés A_1 et A_2 se rencontrent sur une composante X. On a les relations suivantes entre les composantes voisines de X et la composante ambiguë X :

$$c1 \text{ et } c2 \in A_1 \text{ et } c3 \in A_2$$
$$X = \text{voisin}_1^+(c1),\ c1 = \text{voisin}_1^-(X)$$
$$X = \text{voisin}_1^-(c2),\ c2 = \text{voisin}_1^+(X)$$
$$X = \text{voisin}_1^+(c3)$$

Un ensemble de règles permet de traiter cette configuration. Les règles appellent des procédures image qui recherchent les *seconds* voisins des composantes $c1$ et $c3$. Les seconds voisins, sont les premiers rencontrés après la composante X qui est leur premier voisin. Ils sont notés respectivement y et z :

$$y = \text{voisin}_2^+(c1)$$
$$z = \text{voisin}_2^+(c3)$$

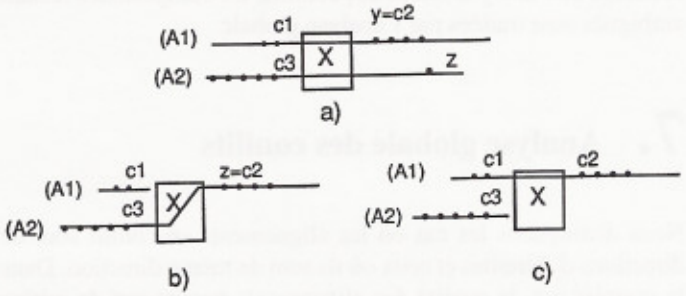

Fig. 7. – Différents chemins possibles pour A1 et A2 suivant la position des seconds voisins y et z.

Les cas où les alignements peuvent être modifiés de manière sûre sont décrits par les règles suivantes :

Règle R1

SI $X = \text{voisin}_1^+(c1)$
$X = \text{voisin}_1^+(c3)$
$y = \text{voisin}_2^+(c1) = c2 \in A_1$
$z = \text{voisin}_2^+(c3) \neq \phi$ et $z \notin A_1$

ALORS
$X \in A_1, X \in A_2$ et $z \in A_2$

La composante X reste affectée aux deux alignements à ce stade de l'analyse et on poursuit l'alignement 2 sur la composante z [figure 7–a].

Règle R2

SI $X = \text{voisin}_1^+(c1)$
$X = \text{voisin}_1^+(c3)$
$z = \text{voisin}_2^+(c3) = c2 \in A_1$

ALORS
$X \in A_2, c2 \in A_2$ et toutes les composantes A_1 à droite de $c2$ sont affectées à A_2

X n'est plus ambiguë et appartient à l'alignement 2 seul [figure 7–b]. Cette règle corrige les situations où les parties de deux alignements proches ont été groupées par erreur, par exemple lorsque du texte est inscrit dans l'interligne du texte principal. Dans tous les autres cas, on choisit l'alignement 1 pour la composante X [figure 7-c].

Dans les configurations d'embranchements en U, on recherche de même les seconds plus proches voisins des composantes $c1$ et $c3$, notés y et z respectivement. On recherche de plus un voisin k de la composante X situé dans un voisinage plus étendu. Comme précédemment les règles relatives aux embranchements en U s'attachent à continuer les alignements, malgré la présence de la composante ambiguë X qui sera soit conservée dans les deux alignements à ce stade de l'analyse, soit affectée à un alignement unique.

A l'issue de l'analyse locale des conflits, les composantes restant ambiguës sont traitées par l'analyse globale.

7. Analyse globale des conflits

Nous distinguons les cas où les alignements en conflit sont de directions différentes et ceux où ils sont de même direction. Dans le premier cas, la qualité des alignements perçus sert de critère d'affectation de la composante ambiguë. Dans le second cas, deux facteurs liés à l'imbrication des alignements et à l'inclusion de la composante dans chacun des alignements, sont combinés pour définir un critère d'affectation.

7.1. ALIGNEMENTS DE DIRECTIONS OPPOSÉES

Dans la situation de conflit de type croisement, l'analyse globale cherche à éliminer un des alignements en conflit. Celle–ci se fonde sur la qualité, du point de vue perceptif, des alignements. Dans un groupement, les caractéristiques locales sont moins visibles que les caractéristiques globales : la perception du tout détermine celle des parties. L'influence des composantes orientées de façon opposée à la direction globale de l'alignement doit donc être modérée. Par contre le nombre de composantes formant l'alignement et l'encombrement spatial de l'alignement sont deux caractéristiques qui se combinent pour faire apparaître la structure de l'alignement. Pour caractériser quantitativement la structure plus ou moins forte d'un alignement, on calcule un facteur de qualité. Soient NC le nombre de composantes de l'alignement, MD le nombre de points d'ancrage de l'alignement ayant la même orientation que celle de l'alignement, DD le nombre de points d'ancrage ayant une orientation différente. Le facteur de qualité P est :

$$P = P_1 * P_2$$

avec $P_1 = \frac{1}{1+\frac{1}{NC-2+\frac{MD}{NC}}}$ si $NC > 1$

$P_1 = 0$ sinon

et $P_2 = \frac{1}{1+\frac{DD}{NC}}$

On notera que $MD \geq 1$ car chaque alignement contient son point d'ancrage origine. Un alignement réduit à une composante unique aura un facteur de qualité P nul.

Si on suppose que toutes les composantes d'un alignement de direction donnée sont des points d'ancrage de même direction, PDD sera nul et P évoluera entre 0.5 et 1 (à la limite) suivant le nombre de composantes de l'alignement.

Dans le cas du croisement de deux alignements, la composante ambiguë appartiendra à l'alignement de meilleure qualité : celui–ci sera conservé, le deuxième éliminé. La différence de qualité doit cependant être suffisamment conséquente.

7.2. ALIGNEMENTS DE MÊME DIRECTION

Dans les configurations de type embranchements, l'analyse globale va permettre soit d'affecter de façon univoque la composante ambiguë à un alignement, soit de constater le chevauchement et de couper la composante en deux parties. L'information contextuelle relative à la position des lignes de référence de l'écriture est utilisée lors de cette analyse.

7.2.1. *Localisation directe des zones médianes de l'écriture*

Une première méthode consiste à localiser sur la composante ambiguë X les zones médianes de l'écriture correspondant à chacun des deux alignements en conflit. Ceci n'est possible que si X résulte effectivement du chevauchement de deux lignes et si chacune des unités d'écriture de la composante X est suffisamment dense dans la zone médiane. Si ces deux conditions sont remplies, le profil des projections de la composante X sur l'axe vertical donne deux pics que l'on peut détecter par une procédure décrite dans [14]. Sur la figure 8, on distingue deux maxima locaux et les pics associés. La position et la largeur des pics donne la position et la hauteur des zones médianes de l'écriture correspondant à chacune des unités d'écriture.

Une fois les zones médianes trouvées, on découpe la composante X en deux parties. La position du point de coupure correspond au minimum du profil de projection entre les deux pics.

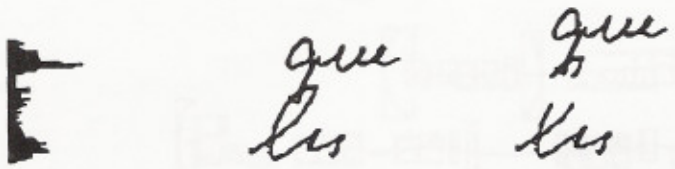

Fig. 8. – Une composante connexe ambiguë, son profil de projection et sa séparation en deux parties.

7.2.2. *Configuration des alignements*

Quand le profil de projection de la composante ne fait pas apparaître de vallée, la composante reste ambiguë. Une deuxième méthode consiste à connaître la position relative des alignements et à tester la présence, au sens d'un certain critère, de la composante ambiguë dans chaque alignement.

Si deux alignements A_1 et A_2 de même direction ont une zone en commun [figure 9] de hauteur L_c, on définit les *degrés d'imbrication* sep_1 et sep_2 des alignements par :

$$\text{sep}_i = \frac{L_c}{H(A_i)} \quad i = 1, 2$$

Les positions extrêmes de chaque alignement, sont calculées en considérant les composantes de l'alignement au voisinage de la composante ambiguë. Les alignements sont dits séparés si la double condition sep_1 et sep_2 faibles est vérifiée.

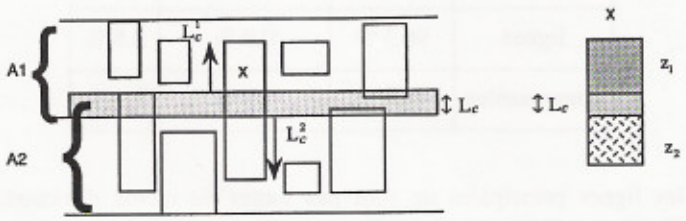

Fig. 9. – Une composante ambiguë x est incluse dans les alignements A_1 et A_2. La zone commune aux deux alignements est en grisé.

On distingue trois parties dans la boîte englobante de la composante X. Une partie z_1 (respectivement z_2) située entre les positions extrêmes de l'alignement A_1 (respectivement A_2) à l'exclusion de la zone commune, et une partie commune aux deux alignements.

Si les alignement sont suffisamment séparés, le *degré d'inclusion* d_inclusion_i de la composante X dans chaque alignement A_i est calculé par la hauteur relative L_c^i de la composante incluse dans la zone de l'alignement.

$$\text{d\_inclusion}_i = \frac{L_c^i}{H(A_i)} \quad i = 1, 2 \qquad (1)$$

Si les alignements ne sont pas nettement séparés, les zones réservées à chaque alignement ne sont pas distinctes et le degré d'inclusion de la composante risque d'être sensiblement le même pour les deux alignements. Le degré d'inclusion est alors la densité relative de la composante X dans l'alignement A_i. Si la densité d'information est suffisante, la composante appartient à l'alignement.

$$\text{d\_inclusion}_i = \frac{P(X, z_i)}{P(X, S_X)} \quad i = 1, 2 \qquad (2)$$

En fonction de la séparation plus ou moins nette des alignements en conflit, les degrés d'inclusion correspondants à chacun des alignements sont calculés avec l'une des formules (1) ou (2). Si un des degrés d'inclusion a une valeur élevée, la composante appartient à l'alignement concerné et on la retire de l'autre. Si les deux degrés d'inclusion ont des valeurs élevées, la composante appartient aux deux alignements : on devra la couper en deux, chacune des parties étant attribuée à l'alignement concerné. On peut envisager un découpage optimal respectant la morphologie des unités graphiques, mais cela n'est pas développé ici. Si aucun des degrés d'inclusion n'a de valeur élevée, aucune décision n'est prise mais l'information d'ambiguïté est conservée en vue de traitements ultérieurs.

8. Résultats et conclusion

Nous avons travaillé sur des images provenant d'enveloppes postales et de brouillons. Les enveloppes offrent un cadre intéressant de test de notre méthode car l'écriture est non contrainte et génère de nombreux cas de conflits. La base de test est composée de 56 documents, dont 35 sont des images de blocs adresse.

Nous donnons les résultats quantitatifs séparément pour les documents de type enveloppes et ceux de type textes. Nous considérons qu'une ligne est détectée quand les composantes de la ligne d'écriture ont été affectées à un alignement, correspondant à cette ligne. La ligne peut cependant être incomplète : composantes restant non affectées ou signes diacritiques affectés à un alignement voisin. Une ligne est non détectée quand ses composantes ne sont affectées à aucun alignement. Une ligne est détectée de manière erronnée quand un ensemble des composantes de cette ligne, appartient à un alignement qui correspond à plusieurs lignes d'écriture. Dans les documents de type texte, une ligne est détectée de manière fragmentée quand à cette ligne correspondent plusieurs alignements qui mis bout à bout reconstituent la ligne d'écriture.

a) Enveloppes

L'écriture des adresses postales est moins cursive que celle des textes provenant de lettres ou de brouillons. Cela est dû à la présence des chiffres, majuscules et abréviations. Les points d'ancrage y sont donc plus rares, phénomène amplifié par le fait que les lignes sont courtes. Une solution consiste à forcer la détection des points d'ancrage dans la direction horizontale. La figure 10 montre le résultat de la segmentation en lignes sur une image de bloc adresse. Les deux premières lignes, collées, ont été séparées.

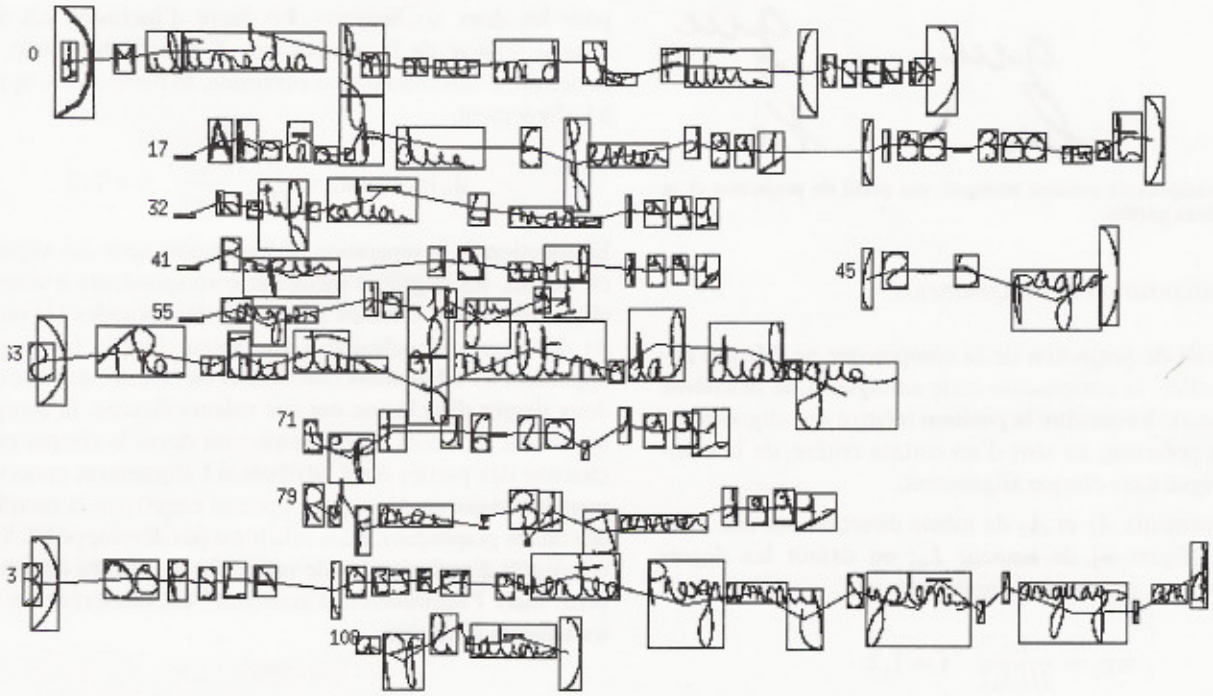

Fig. 11. – Segmentation d'un brouillon.

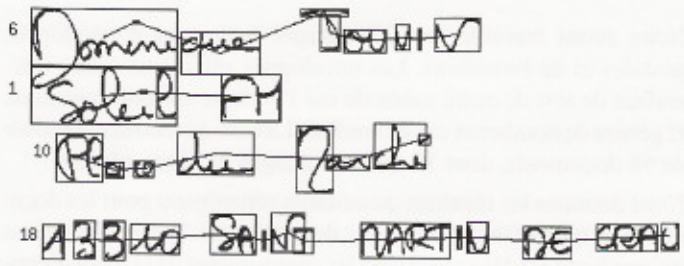

Fig. 10. – Segmentation d'un bloc adresse.

Sur les enveloppes analysées, le taux d'erreur de détection de ligne est de 2,8 %. En effet, seules trois lignes sur la même enveloppe ont été mal détectées. Quelques lignes n'ont pas été détectées ou sont restées incomplètes. En effet, les alignements de trop faible qualité ne sont pas considérés comme des lignes de texte et sont détruits. Ceci arrive sur des lignes courtes composées seulement d'une ou deux composantes (mots cursifs). La qualité d'un alignement devrait faire intervenir son encombrement spatial dans la page.

Une ligne est incomplète quand certaines de ses composantes telles que des accents ou des fragments de caractères restent isolés ou sont attribués à une ligne voisine. Plus précisement, les taux de détection pour les composantes connexes et les lignes des enveloppes sont présentés dans le tableau I.

b) Textes

Les pages de documents ont une structure plus diverse que celle des enveloppes. Un texte est composé de lignes principales, auxquelles peuvent s'ajouter des lignes secondaires telles que insertions, collées ou non aux lignes principales. D'autre part,

Tableau 1. – Enveloppes postales.

	Détection	Non détection	Erreur
lignes	96,3 %	0,9 %	2,8 %
composantes	97,6 %	1,3 %	1,1 %

les lignes principales ne sont pas toutes de même direction. Des composantes non textuelles sont souvent présentes : ratures, encerclements, soulignements, marques d'insertions.

Dans les expériences que nous avons menées, quand le texte ne comportait que des lignes horizontales, nous avons imposé la direction horizontale comme pour les enveloppes.

Vu le nombre élevé de composantes connexes dans ce type d'image, nous donnons seulement les résultats pour les lignes dans le tableau II. Plusieurs lignes ont été détectées en plusieurs alignements, qui mis bout à bout reconstituent des lignes véritables. Ceci est dû à la présence de nombreux points diacritiques. Le taux de lignes trouvées fragmentées est de 7.5 %. D'autres lignes n'ont pas été détectées, ou ont été détectées partiellement car des alignements formant tout ou en partie la ligne ont pu être détruits à une étape du traitement. Les erreurs sont souvent le fait de lignes en biais qui ont une direction intermédiaire à celle des directions choisies pour déterminer les orientations des points d'ancrage. D'autres directions de points d'ancrage devront être créées.

La méthode est sensible au caractère fragmenté de l'écriture, ainsi qu'aux écritures contenant des caractères surplombant d'autres

Tableau 2. – Textes.

	Détection	Détection fragmentée	Non détection	Erreur
lignes	82 %	7,5 %	4,3 %	5,7 %

composantes. Ceci crée des composantes ayant des composantes internes, ce qui gêne la détection des lignes. Un prétraitement futur consistera à repérer les composantes connexes complexes, contenant d'autres composantes ou de taille anormalement élevée. Ces composantes seront retirées dans un premier temps, puis ré–insérées.
D'autre part, les conflits sont d'autant mieux réglés que les composantes chevauchantes sont à l'intérieur des alignements et non en bordure.

La figure 11 est le résultat de la segmentation sur un texte. Chaque alignement est repéré par un identificateur placé en début de ligne. Un trait rejoint les centres de gravité des composantes connexes appartenant au même alignement. Plusieurs lignes, collées initialement, ont été séparées.

La procédure de segmentation a été réalisée en langage C. Le temps de traitement d'une enveloppe est d'environ 3 secondes de temps CPU sur station Sun SPARC–5. Il est de l'ordre de 11 secondes pour une page de texte numérisée à la résolution de 100 dpi.

Nous avons proposé une méthode de résolution de conflits d'alignements lors d'un processus de groupement. Lorsque les alignements sont de même direction, le choix d'un des alignements pour une composante connexe, est celui qui assure la meilleure continuité, au sens de l'organisation perceptive. Cette continuité est vérifiée soit à partir de la recherche de prolongements lors de l'analyse locale des embranchements, soit à partir de la position relative de la composante dans les alignements lors de l'analyse globale. Lorsque les alignements sont de directions différentes, le choix d'un des alignements pour la composante au carrefour des deux, est basé sur la qualité perceptive des alignements. Les améliorations futures consistent à repérer les composantes complexes, constituer de nouveaux alignements à partir des composantes isolées et créer de nouvelles directions de points d'ancrage pour les orientations en biais.

Remerciements : nous remercions Michel Guilloux et le SRTP pour nous avoir fourni des données de test, ainsi que les rapporteurs de cet article pour leurs remarques constructives.

BIBLIOGRAPHIE

[1] G. Nagy, S. Seth, « Hierarchical Representation of optically scanned documents », *7th Int. Conf. on Pattern Recognition*, Montreal, 1984, pp. 347-349.
[2] K. Wong, R. Casey, F. Wahl, Document analysis system, *I.B.M. Journal of Research and Development*, 26, n°6, 1982.
[3] L.A. Fletcher, R. Karturi, Text string segmentation from mixed text/graphics images, *IEEE PAMI*, Vol. 10, N°3, 1988, pp. 910-918.
[4] E. Meynieux, S. Seisen, K. Tombre, « Bilevel Information Recognition and Coding in Office Paper Documents », *8th Int. Conf. on Pattern Recognition*, Rome, 1989, pp. 442-445.
[5] C. Viard–Gaudin, D. Barba, « Extraction robuste et structuration des informations par une approche multirésolution pour la localisation du bloc adresse sur les objets postaux plats », *Actes de CNED'92, Bigre n° 80*, 1992, p. 48-56.
[6] V. Shapiro, G. Gluhchev G., V. Sgurev, Handwritten document image segmentation and analysis, *Pattern Recognition Letters*, N°14, 1993, pp. 71-78.
[7] E. Cohen E., J. Hull, S. Srihari, Understanding handwritten text in a structured environment : determining zip codes from addresses, *Int. Journal of Pattern Recognition and AI*, Vol. 5, N°1 & 2, June, World Scientific, 1991, pp. 221-264.
[8] Th. Paquet, R. Mullot, R. Trupin, K. Roméo, Y. Lecourtier, Un algorithme rapide de détection des mots d'un texte manuscrit, *Congrès AFCET-RFIA*, 1989, Paris, p. 1501-1510.
[9] V. Govindaraju, R. Srihari, S. Srihari, Handwritten text recognition, *Actes de Document Analysis Systems DAS 94*, Kaiserlautern, Octobre, 1994, pp. 157-171.
[10] A.C. Downton, C. Leedham, Preprocessing and presorting of envelope images for automatic sorting using OCR, *Pattern Recognition*, Vol. 23, N°3–4, 1990, pp. 347-362.
[11] L. Likforman–Sulem, C. Faure, Extracting text lines in handwritten documents by perceptual grouping, in Advances in handwriting and drawing : a multidisciplinary approach (C. Faure, P. Keuss, G. Lorette, A. Winter, eds), Europia, 1994.
[12] P. Buser, M. Imbert, *Vision*, Herman, 1987, pp. 404-436.
[13] P. Guillaume, *La psychologie de la Forme*, Flammarion, 1979, chapitre III.
[14] A. Wang,« Détection des lignes dans les textes manuscrits non contraints », *Rapport de stage de* DEA, DEA IARFA, ParisVI, 1994.

Manuscrit reçu le 1er Février 1995.

LES AUTEURS

Laurence LIKFORMAN–SULEM

Laurence Likforman–Sulem est ingénieur des Télécommunications et titulaire d'une thèse de Doctorat en Automatique et Traitement du Signal de l'École Nationale Supérieure des Télécommunications (1989). Elle est Maître de Conférence à l'ENST depuis 1993, ses recherches étant consacrées au traitement de l'écrit. Après avoir travaillé sur des aspects liés à l'authentification, elle s'intéresse actuellement à l'analyse de documents appliquée à la page manuscrite.
Laurence Likforman-Sulem est membre actif du Groupe de Recherche en Communication Ecrite lié à l'Afcet, membre de l'International Graphonomics Society et de la Pattern Recognition Society.

Claudie FAURE

Après une maîtrise de physique à l'Université de Nice, Claudie Faure s'est orientée vers le traitement du signal et la reconnaissance des formes, d'abord à l'université d'Orsay puis à l'université de technologie de Compiègne où elle a obtenu une thèse d'État en 1982. Elle est chargée de recherche au CNRS et travaille depuis 1985 dans l'URA CNRS 820 à l'ENST. Sa recherche actuelle concerne les modes d'expression graphique du point de vue de la perception, de la production, de leur traitement informatique et de leur usage dans des systèmes interactifs.