

Un graphe de voisinage basé sur l'utilisation des distances discrètes

Ivan Leplumey, Charles Queguiner

▶ **To cite this version:**

Ivan Leplumey, Charles Queguiner. Un graphe de voisinage basé sur l'utilisation des distances discrètes. CIFED 2002 - Colloque international francophone sur l'ecrit et le document, Oct 2002, Hammamet, Tunisie. hal-00763418

HAL Id: hal-00763418
https://inria.hal.science/hal-00763418

Submitted on 12 Dec 2012

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.

Un graphe de voisinage basé sur l'utilisation des distances discrètes

Ivan Leplumey * — **Charles Quéguiner****

* *INSA-IRISA*
20, av. des Buttes de Coesmes - 35043 Rennes Cedex, France
** *IFSIC-IRISA*
Campus Universitaire de Beaulieu - 35042 Rennes Cedex, France
{Ivan.Leplumey,Charles.Quéguiner}@irisa.fr

* *INSA de Rennes,* ** *IFSIC - Université de Rennes I*

RÉSUMÉ. *Les distances discrètes sont largement utilisées en traitement d'images : morphologie mathématique, extraction rapide de squelette, reconstruction... Après une introduction, cet article présente dans une première partie (section 2), le principe de construction d'un graphe de voisinage entre les composantes connexes d'une image. Ce graphe, basé sur l'utilisation des distances discrètes, établit une relation entre les points contours de chacune des composantes, pour permettre une meilleure connaissance des positionnements relatifs des unes par rapport aux autres. Cette première partie s'appuie sur une implémentation classique pour la construction du graphe. La seconde partie (section 3) de l'article propose une implémentation particulière qui permet une économie mémoire importante tout en extrayant un graphe qui bien que non équivalent, porte des informations de même nature (voisinage et distance). Enfin, nous conclurons par un bilan comparatif entre les implémentations et des perspectives d'utilisation d'un tel graphe.*

ABSTRACT. *The discrete distances are largely used in image processing : mathematical morphology, fast extraction of skeleton, rebuilding... After an introduction, this paper presents in a first part (section 2), the principle of construction of a graph of vicinity between related components. This graph, based on the use of the discrete distances, establishes a relation between the border points of each component, to allow a better knowledge of relative positionings of those components. This first part is based on a traditional implementation for the construction of the graph. The second part (section 3) of the paper proposes a particular implementation which allows a significant memory economy while extracting a graph. This graph although nonequivalent to section 2 graph, carries identical information (vicinity and distance). This graph makes it possible to answer the problem of relative positioning of the related components. Lastly, we will conclude by comparative assessment and some prospects.*

MOTS-CLÉS : *Pavage de Voronoï, distance discrète, graphe de voisinage, structuration de document.*

KEYWORDS: *Paving of Voronoï, discrete distance, vicinity graph, structuring of document.*

1. Introduction

La compréhension du contenu (tableaux, formules mathématiques, schémas...) de certains documents nécessite, non seulement une classique segmentation en composantes connexes, mais également la prise en compte de relations de voisinage entre les différentes composantes connexes de l'image de chaque document. Ainsi, les relations spatiales (à côté, au-dessus...) entre les signes d'une formule mathématique ou d'une partition musicale sont importantes pour arriver à interpréter l'image binaire du document.

Le fait de qualifier au mieux la relation de voisinage entre deux composantes nous semble un élément primordial sur lequel on peut s'appuyer pour interpréter une image. La qualification proposée consiste à obtenir les liens entre deux composantes proches **A** et **B**. Ces liens sont représentés par un doublet de points **a (x1,y1)** et **b (x2,y2)**, où le point **a** (resp.**b**) appartient au contour de la composante **A** (resp. **B**). Ayant défini une mesure de distance, ces deux points **a** et **b** sont choisis de telle sorte que la distance de **a** vers **b** soit minimum. Nous pensons qu'une structure de graphe est la structure la plus appropriée pour gérer un tel ensemble le liens.

Nous développerons donc par la suite, une méthode basée sur des techniques connues pour extraire un tel graphe de voisinage.

2. Construction du graphe de voisinage

La construction du graphe se découpe en trois phases (cf Figure 1) :

– calcul des composantes connexes ;

– calcul d'une image de distances avec propagation conjointes d'attributs ;

– calcul du graphe.

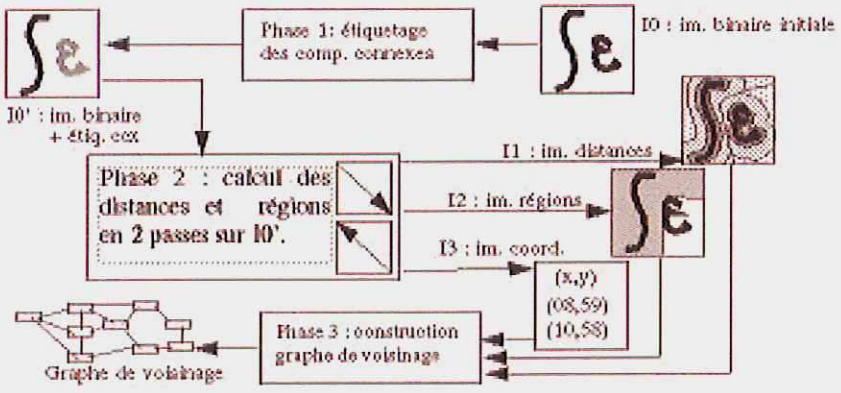


Figure 1: Schéma général.

2.1 Composantes connexes (phase 1)

L'image d'origine est binaire et constituée d'objets posés sur un fond(cf Fig. 2 A). La phase initiale consiste à calculer les composantes connexes de l'image, Le calcul s'appuie sur un algorithme classique, la structure décrivant l'image d'origine étant soit de nature matricielle, soit sous forme de listes d'empans. La structure résultat est de même nature que la structure d'origine.

2.2 Image de distances et attributs (phase 2)

La seconde phase s'appuie sur l'algorithme de Rosenfeld et Pfaltz [ROS 66] et sur le choix d'une distance discrète (distance **d4** du "**city block**", distance **d8** du "**chess-board**", distance de **chanfrein 2-3, 3-4, 5-7** ou autre). Ce choix va influer sur les résultats obtenus, mais n'a jamais été critique dans les applications que nous avons traitées. En utilisant la distance choisie, cette phase va effectuer le calcul, pour tout point du fond, de sa distance à la composante la plus proche (cf Figure 2 B)

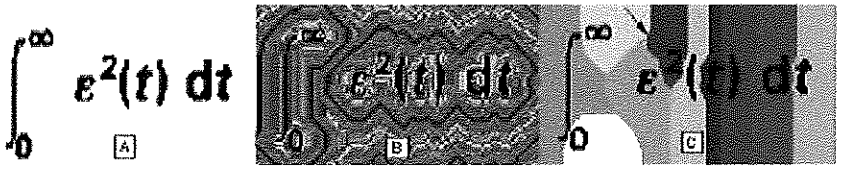


Figure 2: **A**-Image originale (formule mathématique). **B**- Image des distances (d4 : city block) entre les composantes connexes. **C**- Pavage de Voronoï de l'image d'origine en appliquant la distance discrète d4.

L'algorithme cité travaille en deux passes sur l'image des composantes connexes du document ; il nécessite la mémorisation d'une image intermédiaire (**I1**) de la taille de l'image d'origine (**I0**). Dans cette image **I1**, les pixels (x, y) ont une valeur associée :

– zéro ; pour les pixels appartenant à des composantes (dans notre exemple, les caractères imprimés de la formule mathématique)

– entier(>0) pour les pixels appartenant au fond ; cet entier est la distance à la composante la plus proche.

En même temps que ce calcul de distance, nous effectuons dans une seconde image (**I2**), la propagation conjointe du nom de la composante la plus proche. Ceci nous donne pour l'image d'origine (**I0**), le pavage en polygones de Voronoï [BOR 86], [MEL 92] ; pavage qui correspond à la juxtaposition des zones d'influence de chaque composante connexe de l'image étudiée (cf Figure 2 C).

2.3 Extraction d'un graphe de voisinage (phase 3)

Tout point du plan est étiqueté (dans l'image **I2**) par la phase précédente, avec le nom de la composante la plus proche (polygones de Voronoï). Il est alors possible par un simple parcours de l'image **I2** d'établir les relations entre les différentes régions connexes. Ainsi sur la figure 3, le symbole "**infini**" et l'exposant "**2**" sont en relation, leurs régions associées étant connexes. Nous obtenons ainsi une première relation de voisinage entre les régions, ce qui correspond à un graphe de Delaunay (Fig 3 A).

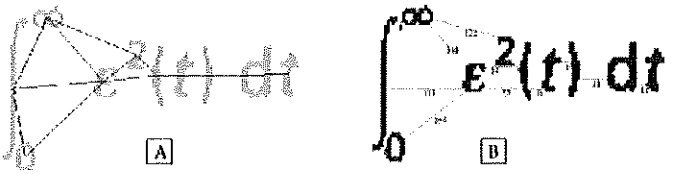


Figure 3: **A**- Graphe de Delaunay (lien entre composantes voisines). **B**- Graphe entre les points contours des différentes composantes.

De plus, nous pouvons directement valuer les arcs de ce graphe par une distance entre les composantes "voisines" en effectuant une simple consultation dans l'image

des distances (**I1**). On conserve seulement la distance minimum observée entre deux composantes, cette valeur étant la plus informante.

2.4 Enrichissement du graphe

Il est possible d'enrichir ce graphe par la connaissance des points contours ayant conduit à lier deux composantes. Pour ce faire, on propage de nouveaux attributs dans une image (**I3**), conjointement à la distance (image **I1**) et au nom de la composante (image **I2**) : les coordonnées (xOrigine, yOrigine) du point depuis lequelle s'est propagée la distance au contour de la composante (cf Fig. 3 B).

L'exploitation de ces nouveaux attributs permet d'accroitre la connaissance sur le lien le plus proche qui lie deux composantes. Dans un contexte d'interprétation d'images, ce graphe peut se révéler d'une grande utilité.

2.5 Critique des résultats obtenus

En terme d'implémentation, les besoins en espace mémoire sont très importants. Ainsi, on doit en sus de l'image binaire d'origine (**I0**), manipuler une structure occupant un espace équivalent à quatre images supplémentaires pour mémoriser les résultats de propagation (une pour la distance, une pour les noms et deux pour les coordonnées xOrigine et yOrigine des points contour).

Au niveau de la pertinence des arcs du graphe et de leur valuation, si l'on regarde la connexion entre le symbole "**infini**" et l'exposant "**2**" en observant le pavage, on note que les deux régions se rencontrent au dessus du "**e**" et que le lien calculé n'est donc pas toujours celui de plus court chemin. On peut ainsi voir sur le schéma suivant ("flèche" sur la Figure 4 A et sa correspondante sur la Figure 3 B), le lien entre le symbole "**infini**" et l'exposant"**2**" qui se décompose en deux segments joints à l'endroit où se rencontrent les deux régions. En cas d'égalité de distance entre deux régions, seul le premier chemin trouvé est conservé.

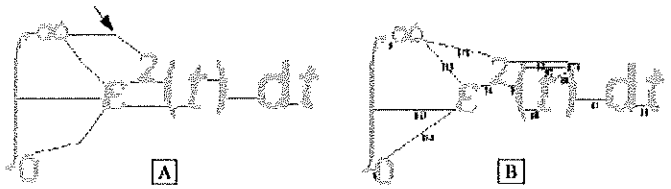


Figure 4: **A**- Décomposition des liaisons (chemin suivi lors de leur création). **B**- Graphe résultat d'une seule passe (NW->SE) sur l'image.

La section suivante propose une implémentation différente de la deuxième phase, ainsi que quelques variantes, permettant finalement d'aboutir plus rapidement à un graphe de voisinage pertinent quoique légèrement différent, tout en limitant les besoins d'espace mémoire.

3. Une implémentation différente

L'algorithme précédent ayant des besoins importants en mémoire centrale, le but de cette section 3 est de montrer que le graphe calculé (phase 2) à l'aide d'une mémorisation locale sur deux lignes (ligne précédente, ligne courante) donne des résultats satisfaisants. Avec l'implémentation décrite à la sous-section 3.1, le graphe de voisinage est construit à la volée en même temps que sont calculés les quatre paramètres précédemment définis : distance, nom, xOrigine, yOrigine. Les phases 2 et 3 sont donc regroupées en une seule (cf Figure 5 Phase B).

Comme pour l'algorithme de Rosenfeld et Pfaltz[ROS 66], il est préférable au cours d'une seconde passe (de bas en haut) d'éliminer les liens de voisinage peu significatifs issus de la première passe (de haut en bas). Cette élimination peut se faire en suivant plusieurs politiques dépendantes des besoins de l'application finale, politiques que nous aborderons dans la sous-section 3.2.

Enfin, nous proposons une amélioration de cette première implémentation dans la sous-section 3.3, dans le but d'améliorer la propagation des noms et des distances, propagation moins bien réalisée, lorsqu'il n'y a pas de mémorisation globale. Le schéma suivant (cf Fig. 5) illustre cette nouvelle implémentation.

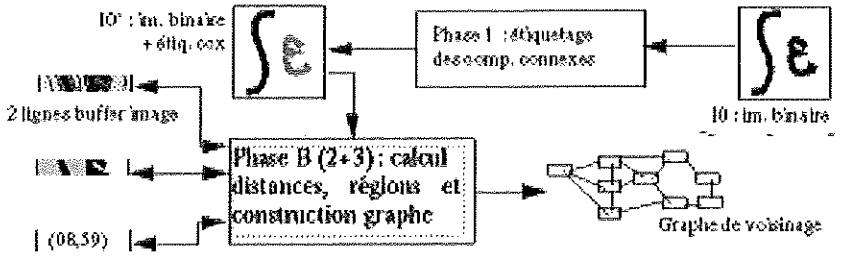


Figure 5: Schéma général modifié (construction du graphe à la volée).

3.1. *Construction du graphe en une passe*

L'application sur l'image origine d'une seule passe du haut vers le bas, de la gauche vers la droite, donne les résultats sur la Figure 4 B. Sur cette image exemple, les deux parenthèses et la lettre "**d**" débutent en haut sur la même ligne de pixels, ce qui permet au chiffre exposant "**2**" de se connecter à ces trois composantes voisines. Sur la Figure 6 sont représentées les régions issues du balayage (haut-bas ou NW->SE) ; ces régions expliquent les liaisons du chiffre exposant "**2**" avec les trois composantes "**( ) d**" (voir les trois flèches : Fig. 6 A).

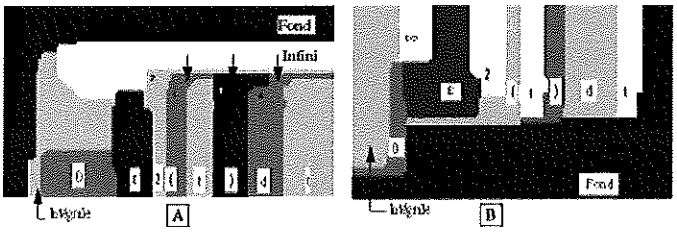


Figure 6: Zones d'influence des composantes connexes :**A**- première passe (NW->SE) et **B**- passe bas-haut (SE->NW) seule.

Le balayage inverse, du bas vers le haut, et de la droite vers la gauche (SE->NW) conduit à des zones d'influences évidemment différentes (Fig. 6 B) et donc à un graphe différent (Fig. 7 A).

On s'aperçoit qu'une coopération des deux graphes (Fig. 3 B et Fig. 7 A) pourrait amener à une conservation des liaisons les plus significatives. Ainsi, sur notre exemple, les liaisons établies entre le chiffre "**2**" et la parenthèse fermante "(" , ainsi

qu'entre ce même chiffre "**2**" et la lettre "**d**" sont très probablement à éliminer au vu de l'équation mathématique, un tel lien n'ayant aucun apport vis à vis de l'interprétation de la formule. Pour ce faire, une solution consiste à faire collaborer les résultats des deux passes de l'algorithme.

3.2. *Règles de synthèse des deux passes*

Cette sous-section propose deux stratégies d'élimination des liaisons les moins significatives en s'appuyant sur la fusion des résultats issus des deux passes sur l'image d'origine.

Règle de double existence

La première stratégie consiste à garder comme valides, les liaisons du premier balayage (passe NW->SE), telles qu'il existe pour le second balayage (passe SE->NW) une liaison entre les mêmes composantes. La liaison peut cependant être déplacée par le second balayage à un endroit différent, minimisant la distance (Fig. 7 B). Ainsi, sur l'exemple, le premier balayage met en contact la partie haute des deux parenthèses, alors que le second balayage met en contact leurs parties basses. Le lien entre les deux parenthèses est donc confirmé.

Règle de confirmation de positionnement

La règle précédente a effectivement éliminé les liaisons les moins pertinentes. Sur notre exemple, nous pouvons cependant nous poser la question de la conservation du lien entre les deux parenthèses. Est-il nécessaire, et quelle serait la règle à appliquer pour conserver ou supprimer un tel lien. Ayant noté que la liaison s'établit à des positions différentes suivant le balayage, nous proposons une règle de conservation d'une liaison :

– (1) la liaison existe pour chacun des deux balayages ;

– (2) la liaison met au moins en contact les deux mêmes points pour chacun des deux balayages.

Le résultat d'une telle règle, appliquée à notre exemple, est donné (Fig. 7 C). Le choix des règles est fonction de l'application et du type d'image traitées.



Figure 7: Graphes obtenus :**A**- par la passe bas-haut(SE->NW) seule. **B**- par la règle de double existence (simplification). **C**- par la règle de confirmation de positionnement.

3.3. *Amélioration de la propagation des noms*

Dans notre exposé, hormis le calcul des composantes connexes qui est classique, deux points n'ont pas encore été abordés :

– l'influence du choix de la distance discrète (distance d4, d8, chanfrein 2-3...) ;

– la politique de propagation des attributs associés aux composantes.

Les paragraphes qui suivent vont conduire à l'obtention de zones d'influence sensiblement différentes pour les composantes connexes d'une image. Sachant que le

graphe est basé directement sur le voisinage entre zones d'influence, le graphe suivra les variations des zones d'influence.

Choix de la distance

Quelle que soit la méthode utilisée, avec ou sans mémorisation de matrices de distances, les résultats sont avant tout liés au choix de la distance discrète. Ainsi, le pavage de Voronoï issu de l'emploi de **d8** est différent de celui calculé avec **d4**, ce qui conduit inévitablement à un graphe différent. Une synthèse sur les distances discrètes utilisables se trouve dans [THI 94].

Politique de propagation des attributs

Le second point est soulevé par l'implémentation sans mémorisation ; dans celle-ci, en effet, il n'existe pas de seconde passe (SE->NW) pour corriger le pavage proposé par la première passe (NW->SE).

Elargissement du masque de propagation

Nous proposons donc de modifier le cône de propagation des informations d'une composante pour essayer d'accroitre, en une seule passe, la zone d'influence de cette composante.

Ainsi, pour la distance **d4**, un point propagera son nom et ses attributs dans toute la partie de l'image située en dessous de lui et à sa droite. Cette influence est représentée dans le masque suivant (Fig. 8 A), qui correspond au premier passage sur l'image. Ce masque montre qu'un point **P** va influencer trois de ses points voisins, son voisin de gauche, dont il est à une distance de 1, son voisin du dessous à la même distance et son voisin diagonal Sud-Est qui est éloigné d'une distance 2 du point **P**.

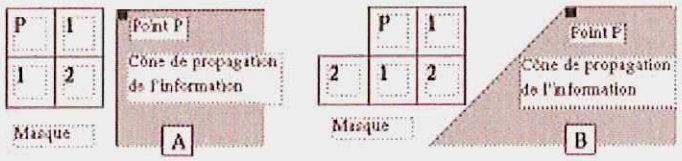


Figure 8: **A**- Masque et cône d'influence pour la distance **d4**. **B**- Masque modifié pour la distance **d4** et cône d'influence induit.

Afin de minimiser l'effet anisotrope induit par le balayage "Droite-Gauche" dans le calcul des distances, nous proposons d'appliquer un autre masque de calcul de distance, qui tout en conservant la distance **d4**, permet un meilleur cône de propagation des informations (Fig. 8 B). Plus, on augmente la taille du masque sur la ligne précédente à droite du point courant (ici, coefficient 2), plus on agrandi l'angle du cône de propagation, mais plus on augmente le coût des calculs de la distance minimum (Fig. 9 A).

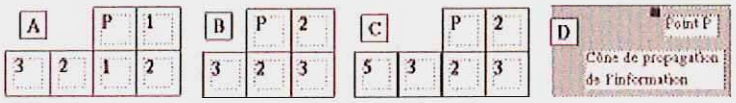


Figure 9: **A**- Masque élargi pour la dist. d4. **B**- Masque d'origine pour chanfrein 2-3. **C**- Masque élargi pour chanfrein 2-3. **D**- Cône d'influence de la propagation étendue.

On peut alors concevoir des familles de masques adaptés aux différentes distances de chanfrein. Ainsi, les figures 8-B et 8-C indiquent les premiers masques de la distance chanfrein 2-3.

Changement de la politique de balayage

Une autre façon d'accroitre la zone d'influence d'un pixel, consiste à adopter une variation dans la politique de balayage de l'image vis à vis de celle proposée par l'algorithme initial de Rosenfeld et Pfaltz [ROS 66]; cette variation peut être perçue comme une décomposition monodimensionnelle, suivant l'horizontale, puis la verticale, de l'algorithme initial bidimensionnel.

Cette politique, appelée propagation étendue, permet d'obtenir la zone d'influence de taille maximum pour un point donné, c'est à dire un demi-plan (Fig. 9 D).

Cette variation de l'algorithme calcule la distance gauche-droite, puis la distance droite-gauche sur la ligne de pixels courante, avant de fusionner ces résultats avec ceux provenant de la ligne de pixels précédente. Le graphe obtenu est différent, l'extension des zones d'influence des composantes se fait de façon symétrique (vers la droite et vers la gauche) par rapport à l'horizontale. On peut observer Fig. 10 A (resp. 10 B) les zones d'influence des différentes composantes de la formule mathématique générées par le balayage Haut-Bas (resp. Bas-Haut).

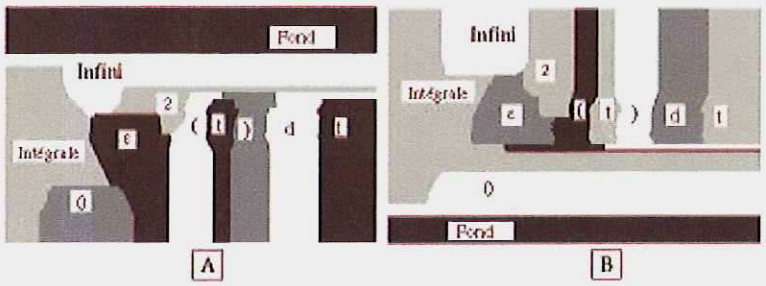


Figure 10: **A**- Propagation étendue (balayage Haut-Bas) : nouvelles zones d'influence des composantes connexes. **B**- Propagation étendue (balayage Bas-Haut).

Notons que la zone horizontale blanche correspond à la propagation du nom associé au symbole "**infini**" ; on voit cette propagation qui s'étend maintenant sur la gauche de l'image (Fig. 10 A), contrairement à l'image précédente (Fig. 6 A). Notons également la zone d'influence de l'exposant"**2**", qui est désormais circonscrite à la partie supérieure de la formule mathématique.

4. Conclusion

Le graphe que nous obtenons va être exploité lors de l'analyse des documents (formule mathématique, partition musicale) en vue d'une reconnaissance de certains éléments structurés composant le document.

Dans cette conclusion, nous donnons quelques éléments de comparaisons entre les deux implémentations, celles des sections 2 et 3, puis nous proposons une perspective d'utilisation de ce graphe à un niveau plus global du document ; ceci en vue d'un pilotage du processus de reconnaissance globale du document.

Comparaison des deux implémentations

La comparaison qualitative est plus délicate car elle est directement liée à l'existence, au positionnement et à la pertinence des liens extraits. Ceux-ci varient en fonction des distances choisies et de l'implémentation. Le principal point noté concerne la sensibilité au sens de balayage de l'image (haut en bas ou bien bas en haut) de la nou-

velle implémentation proposée. L'utilisation des différents graphes dans un processus de traitement de documents permettra d'évaluer si certains graphes sont insuffisants.

Pour une image XxY	Mémorisation globale (cf. section 2)	Construction à la volée (cf. section 3)
Taille mémoire (hors graphe)	4 matrices XxY d'entiers	4 fois 2 lignes de Y entiers (donc X/2 fois moins)
Temps calcul	équivalent	équivalent
Sensibilité au sens de balayage (Haut-Bas ou Bas-Haut)	insensible	sensible

Tableau 1: Comparaison des 2 méthodes de construction du graphe de voisinage.

Perspectives

Le premier but de ce travail concernait l'extraction de liens précis de voisinage dans le cadre de la reconnaissance de documents structurés. L'étude a été appuyée par des résultats sur une formule mathématique où les relations de voisinage spatial ont besoin d'être bien précisées pour permettre une bonne interprétation.

Les paragraphes suivants montrent les utilisations potentielles d'un tel graphe dans l'analyse de documents de nature différente.

Extraction de paramètres de segmentation des régions

Ce paragraphe présente les résultats de notre application sur un texte. On peut compléter le graphe de voisinage, par une analyse statistique des liens, pour extraire du texte les différentes entités le composant (Fig. 11 A).

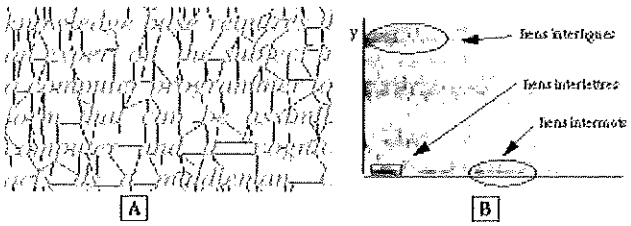


Figure 11: **A**- Propagation étendue : graphe résultat sur un texte (extrait). **B**- Matrice d'occurrences.

A partir du graphe présenté sur l'image précédente, on construit une matrice d'occurrences(Fig. 11 B) qui comptabilise le nombre de vecteurs possédant une certaine direction et une certaine longueur. Ainsi, pour un vecteur allant du point (xOrigine, yOrigine) au point(xFin, yFin), on incrémente une matrice d'occurrences à la position correspondant à l'extrémité du vecteur (lxFin-xOriginel, lyFin-yOriginel).

Sur la matrice d'occurrences est superposé un rectangle près de l'origine. Celui-ci délimite une zone de liens courts et horizontaux; zone qui doit logiquement correspondre aux liaisons entre lettres d'un même mot, ou entre un mot et la ponctuation qui le suit. La véracité de cette hypothèse est illustrée par l'image suivante (Fig. 12 A).

En extrapolant le raisonnement à partir de ces résultats, on peut trouver les liens correspondant aux liens entre la dernière lettre d'un mot et la première lettre du mot suivant sur la même ligne d'écriture (Fig. 12 B).

De même, sur la figure 12 C, on été retenus les liens entre des lettres de deux lignes adjacentes d'écriture.

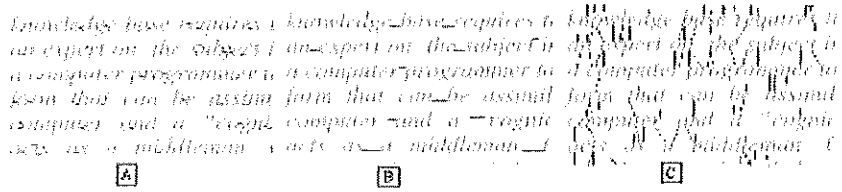


Figure 12: Extrait du texte, avec tracé en noir : A- , les liens interlettres. B- les liens intermots. C- les liens interlignes.

Reconnaissance de documents

Nous développons actuellement un système de reconnaissance de documents structurés que nous avons appliqué en premier lieu aux partitions musicales [COU 95]. Ce système utilise un formalisme grammatical possédant des opérateurs de position qui permettent de définir les positionnements relatifs d'objets les uns par rapport aux autres dans une image. Par compilation d'une grammaire définie à l'aide de ce formalisme, le système génère un analyseur adapté. Actuellement, les opérateurs de position utilisent une distance euclidienne sur des approximations des objets manipulés (par exemple : la boite englobante d'une composante connexe). Le graphe de voisinage sera intégré au système de reconnaissance afin d'offrir dans le formalisme des opérateurs de position beaucoup plus précis. En effet, le graphe de voisinage permet de définir une position en introduisant un contexte graphique et en évitant des approximations puisque celui-ci utilise le contour des objets. L'utilisateur du formalisme pourra alors choisir soit les positionnements existant actuellement, soit ceux définis à l'aide du graphe de voisinage en fonction de la précision souhaitée.

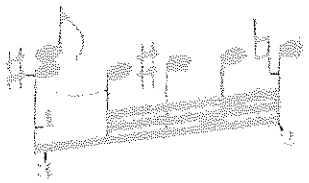


Figure 13: Graphe obtenu sur une partition musicale.

5. Bibliographie

[BOR 86] BORGEFORS G., « Distance transformations in digital images », Computer vision, graphics, and image processing, vol. 34, p. 344-371, 1986.

[COU 95] COÜASNON B.— CAMILLERAPP J., « A Way To Separate Knowledge From Program In Structured Document Analysis : Application To Optical Music Recognition », International Conference on Document Analyse and Recognition (ICDAR'95), p. 1092-1097, 1995.

[MEL 92] MELKEMI M., Approches géométriques par modèles de Voronoï en segmentation d'images Thèse de l'Université Joseph Fourier, Grenoble I, 1992

[ROS 66] ROSENFELD A.— PFALTZ J.L., « Sequential operations in digital picture processing » Journal of ACM,vol. 13(4) p. 471-494, 1966

[THI 94] THIEL E. Les distances de chanfrein en analyses d'images : fondements et applications Thèse de l'Université Joseph Fourier, Grenoble I, 1994