

# Localisation du bloc-adresse sur des objets postaux basée sur la coloration hiérarchique de graphe

Djamel GACEB    Véronique EGLIN    Frank LEBOURGEOIS    Hubert EMPTOZ

LIRIS INSA de Lyon – 20, av. Albert Einstein 69621 VILLEURBANNE CEDEX

{djamel.gaceb1,veronique.eglin,flebourg,hubert.emptoz}@insa-lyon.fr

## Résumé

Un système efficace de tri du courrier postal est principalement basé sur une lecture optique correcte de l'adresse de destination sur l'enveloppe. Cependant, le bloc adresse doit être localisé avant l'opération de reconnaissance par l'OCR. Le fait que cette phase ait un grand impact sur les performances globales du système, elle représente la phase la plus importante du processus de lecture. Pratiquement, une bonne localisation mène à un taux de reconnaissance plus élevé. Statistiquement, la principale cause de rejet des courriers est liée à un échec de l'une des étapes de module de localisation du bloc adresse (LBA). Compte tenu des limites couramment utilisées pour la LBA, dépendantes pour la plupart d'une architecture modulaire linéaire, nous proposons dans cet article une architecture robuste basée sur le principe de la coloration hiérarchique de graphe CHG. Notre stratégie consiste à augmenter la cohérence et la performance des différentes étapes du processus de LBA pour réduire au maximum le rejet des courriers et les temps de traitement.

## Mots clefs

Localisation de texte, segmentation de la structure physique, traitement en temps réel, tri de courriers et de documents, coloration des graphes.

## 1 Introduction

Le tri automatique de courrier postal repose sur une lecture optique des adresses indiquées sur les enveloppes à trier, et leur expédition à la bonne destination. Après l'acquisition de l'image de l'enveloppe par une camera CCD, le système de tri doit pouvoir localiser, en temps réel et avec grande précision le bloc-adresse (BA) et l'envoyer à l'OCR<sup>1</sup>. Dans notre application le BA représente la zone d'intérêt contenant l'information vitale pour reconnaître la destination. Par conséquent, toute

<sup>1</sup>OCR: (Optical Character Recognition) Reconnaissance Optique de Caractères.

adresse mal localisée (et donc mal reconnue) conduit à un rejet immédiat du courrier. Compte tenu du fait que l'adresse de destination est supposée apparaître au coin gauche inférieur pour certains types de courrier, d'autres ne respectant pas cet arrangement strict ne peuvent être convenablement analysés. De plus, la présence de timbres, de marques de la poste, de logos imprimés, d'annonces diverses et autres informations sur le courrier, rend la tâche de localisation plus difficile (Figure 1).



Figure 1 - Présence d'objets parasites autour du bloc adresse

En générale, un module de localisation du bloc adresse (LBA) se compose de trois parties élémentaires: la binarisation et la détection des composantes connexes, la segmentation de la structure physique de l'image de l'enveloppe, et l'extraction du BA (voir Figure 2). Pratiquement, un dysfonctionnement sur l'une de ces étapes réduit les performances des autres, et par conséquent mène à une localisation imprécise de BA ou à une lecture optique erronée de son texte.

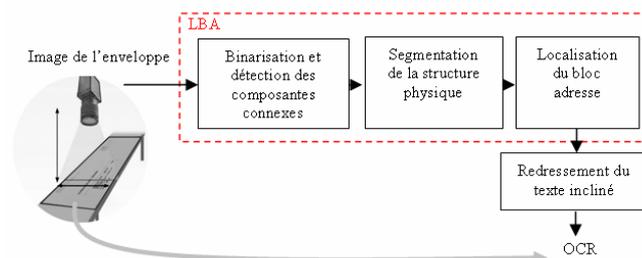


Figure 2 – Etapes élémentaires du module LBA

Durant cette dernière décennie, plusieurs travaux ont été proposés pour améliorer le mieux possible la LBA. D'une manière générale, l'ensemble de ces méthodes peut se diviser en deux grandes classes : celles qui sélectionnent le BA parmi plusieurs blocs candidats et celles qui extraient directement le BA à partir de l'image de l'enveloppe. Les méthodes de la première classe utilisent diverses techniques de segmentation pour faire émerger les différents blocs visiblement existants sur l'image de l'enveloppe. Puis, elles extraient des descripteurs de chaque bloc, pour révéler celui qui contient explicitement l'adresse de destination. Les méthodes de la deuxième classe se limitent à localiser directement le BA sans segmenter l'image de l'enveloppe en plusieurs blocs. Malgré une vitesse de traitement légèrement supérieure à celle des méthodes de la première classe, le taux de mauvaise localisation reste le plus élevé. Dans le cadre de notre étude, nous nous sommes intéressés aux méthodes de la première classe.

En 1988, un système expert a été proposé par Wang et al [1] pour trier automatiquement le courrier. Les auteurs utilisent un tableau noir (blackboard) pour conserver et exploiter les caractéristiques géométriques des blocs obtenus dans la phase de traitement de différents types d'images d'enveloppes. Quelques années après, Viard-Gaudin et Barba dans [2] ont proposé une nouvelle approche de LBA légèrement différente basée sur la construction d'une structure pyramidale de données, dans laquelle, on fait intervenir une analyse descendante pour extraire les relations spatiales entre les divers blocs segmentés avec leurs caractéristiques afin d'en reconnaître la nature. Yu dans [3] utilise presque le même principe adapté cette fois aux courriers complexes. L'approche proposée par Jeong et al [4] est basée sur le groupement des composantes connexes issues de l'image binaire, où chacune d'elle est affectée à une des 9 classes. De ce fait, le bloc de l'adresse de destination est déterminé en ne sélectionnant que quelques classes. Récemment et sans binariser l'image, Eiterer a proposé dans [5] une nouvelle piste, par une approche basée sur les dimensions fractales. Une classification par la méthode K-Means est utilisée pour marquer les pixels en niveaux de gris comme fond, bruit ou objets sémantiques qui constituent les classes de base définissant les timbres, les cachets de la poste, et les blocs d'adresse.

Le but de cet article est de proposer une nouvelle approche de LBA basée sur une coloration hiérarchique de graphe (CHG) et une organisation pyramidale des données. Notre stratégie consiste à augmenter la cohérence et la performance des différentes étapes du processus de LBA pour réduire au maximum le rejet des courriers et les temps de traitement.

## 2 Méthodes existantes

Nous présentons brièvement dans cette partie les différentes techniques de binarisation existantes, de détection des composantes connexes (CCs), de segmentation de la structure physique et de coloration des graphes.

### 2.1 Binarisation et détection des CCs

En général, la binarisation est appliquée dans la première étape de processus de LBA, elle a un impact très fort sur les performances du système de tri. Généralement, les méthodes de binarisation se divisent en deux classes : Globales (exp : méthode d'Otsu [6]), et locales (exp : méthode de Sauvola [7]). On s'aperçoit que les méthodes globales ne peuvent pas fournir de bons résultats quand l'image en niveaux de gris a une luminosité non-uniforme ou un histogramme multimodal. Les méthodes locales dépassent cette limite mais nécessitent plus de calcul et sont ainsi plus lentes. Après l'étape de binarisation, on effectue une analyse des CCs pour extraire les différentes informations vitales aux étapes suivantes. Formellement, une composante connexe (CC) est un ensemble de pixels connectés deux à deux. Dans un BA imprimé, chaque caractère alphanumérique est une CC. Afin de réduire le temps de traitement nécessaire à l'extraction des CCs, plusieurs méthodes ont été développées. Un très bon état de l'art a été effectué dans [9]. Dans notre étude nous nous sommes intéressés au travail de Pavlidis [8] qui a modélisé le problème de détection des CCs par un graphe d'adjacences des lignes (LAG). Cette technique est principalement basée sur la représentation des longueurs des séquences de pixels noirs.

### 2.2 Segmentation de la structure physique

D'une manière générale, la segmentation de la structure physique de l'image de l'enveloppe désigne sa séparation en ses éléments constitutifs. A ce jour, ils existent trois stratégies de segmentations : Descendante (top-down) [12], Ascendantes (bottom-up) [13] [10] et mixte [12]. Déforges et Barba dans [10] ont présenté une méthode générique basée sur une description en multi-résolution de l'image de document utilisée pour la LBA. Une structure presque similaire a été utilisée par Wang [11] pour distinguer les blocs de textes des blocs de graphiques, et les représenter dans un modèle structurel. Shi et Govindaraju dans [12] ont proposé un algorithme basé sur l'application de « fuzzy directional runlength ».

Il est évident que les blocs issus des méthodes de segmentation classiques ne peuvent pas être toujours purs si la zone d'adresse est entourée d'objets parasites (logos, annonces publicitaires, graphiques). Pour obtenir plus de robustesse face à ces limites, nous présentons dans la section 3.2 notre nouvelle approche hiérarchique basée sur le regroupement progressif des CCs en utilisant la coloration hiérarchique de graphe.

### 2.3 Application de la coloration des graphes

Différents problèmes pratiques de classification peuvent être modélisés par la coloration des graphes. La forme générale de ces applications nécessite la formation d'un graphe par les nœuds (sommets) qui représentent les objets d'intérêt et les arcs qui définissent les relations entre ces objets. On veut par exemple décomposer un ensemble d'individus en plusieurs classes homogènes sans savoir a priori leur nombre. Pour cela, il suffit de représenter chaque individu  $i$  par un sommet

vi et d'ajouter un arc  $E(v_i, v_j)$  entre chaque paire d'individus suffisamment différents. La coloration du graphe  $G(V, E)$  consiste à affecter à tous ses sommets une couleur de telle sorte que deux sommets adjacents ne portent pas la même couleur. Le nombre de couleurs utilisées pour colorer le graphe est appelé nombre chromatique, et ces couleurs vont correspondre aux différentes classes d'individus (Figure 3).

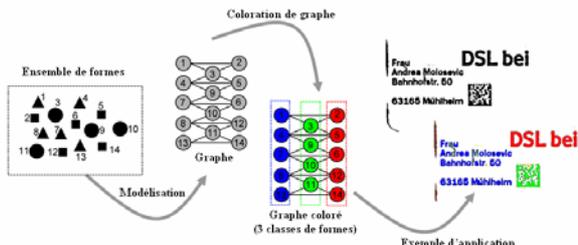


Figure 3 – Application de la coloration des graphes

La coloration est appelée b-coloration, si pour chaque couleur  $c_i$ , il existe au moins un sommet  $v_i$  coloré  $c_i$ , dont le voisinage est coloré par toutes les autres couleurs. Le sommet  $v_i$  est dit un sommet dominant (b-chromatique) pour la couleur  $c_i$ . Des nouveaux algorithmes de coloration et de b-coloration des graphes ont été proposés par Effantin et Kheddouci [14] [17]. Plus de détails sur l'approximation du nombre b-chromatique ont été proposés par Corteel dans [16]. Toutes ces algorithmes ont été efficacement introduits dans les travaux d'Elghazel [15] qui propose une nouvelle méthode non supervisée de classification des données médicales basée sur la b-coloration de graphe où le nombre des classes n'est pas connu à l'avance. Sur la même base de données, la comparaison de la justesse de cette méthode avec celle de la méthode de classification hiérarchique agglomérative, de l'approche du Hansen et de la classification de DRG, a montré que cette technique offre une vraie représentation des classes par les individus dominants et garantit une meilleure disparité interclasses.

Dans ce qui va suivre, nous présentons notre nouvelle approche de LBA. La coloration de graphe est introduite pour corriger la sur ( et/ou la sous) segmentation de la structure physique en blocs liée à la présence des objets parasites autour des zones d'intérêts, et la b-coloration est utilisée pour entraîner un classifieur à sélectionner le bloc-adresse parmi plusieurs candidats.

### 3 Notre nouvelle approche de LBA

Nous présentons dans cette section les différentes étapes de notre approche de LBA. Le schéma de la figure 4 représente notre architecture pyramidale qui permet d'obtenir une grande cohérence entre les différents modules.

#### 3.1 Seuillage et détection des CCs utilisées

La séparation entre l'étape de binarisation et celle de la localisation des zones de texte, augmente à la fois, le temps du calcul et conduit à une sur-segmentation du bruit et de la texture de papier sur des zones vides de l'image. En effet, aucune des méthodes classiques (globales ou locales) ne remplit efficacement toutes les conditions imposées, notamment, une certaine efficacité sur toutes les images pendant une durée de calcul limitée. Nous avons pu optimiser cette étape en appliquant un seuillage local uniquement à proximité des zones de texte (Figure 5) que nous avons localisées par la méthode des gradients cumulés avec la multirésolution et la morphologie mathématique [18].

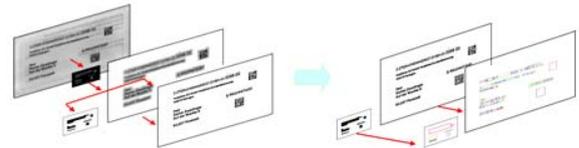


Figure 5 – Seuillage mixte et détection des CCs

Nous détectons ainsi les CCs, de premier plan de la carte binaire et du masque des blocs. La méthode utilisée est inspirée des études de Pavlidis sur la structure LAG (Line Adjacency Graph) [8], structure particulièrement adaptée au balayage ligne par ligne de l'image. Elle consiste à mettre en relation les séquences de pixels noirs de deux lignes consécutives de la fenêtre d'analyse. Chaque cc est représenté par les coordonnées de sa boîte englobante avec :  $cc(i) = (x_i^1, y_i^1, x_i^2, y_i^2)$ . Soit  $V(L_1) / (V(L_3))$  l'ensemble des cc de la carte binaire (/masque des blocs) qui représente le niveau  $L_1 / (L_3)$  le plus fin (/le plus grossier) de la pyramide. Avec  $V(L_k) = \{CC_k(i) / i=1 \dots N_k\}$ ,  $N_k$  c'est le nombre des CCs dans la couche  $k/k=1,2,3$ . Afin de réduire le nombre  $N_k$  nous supprimons tous les CCs parasites. La segmentation de la structure physique va donc se baser

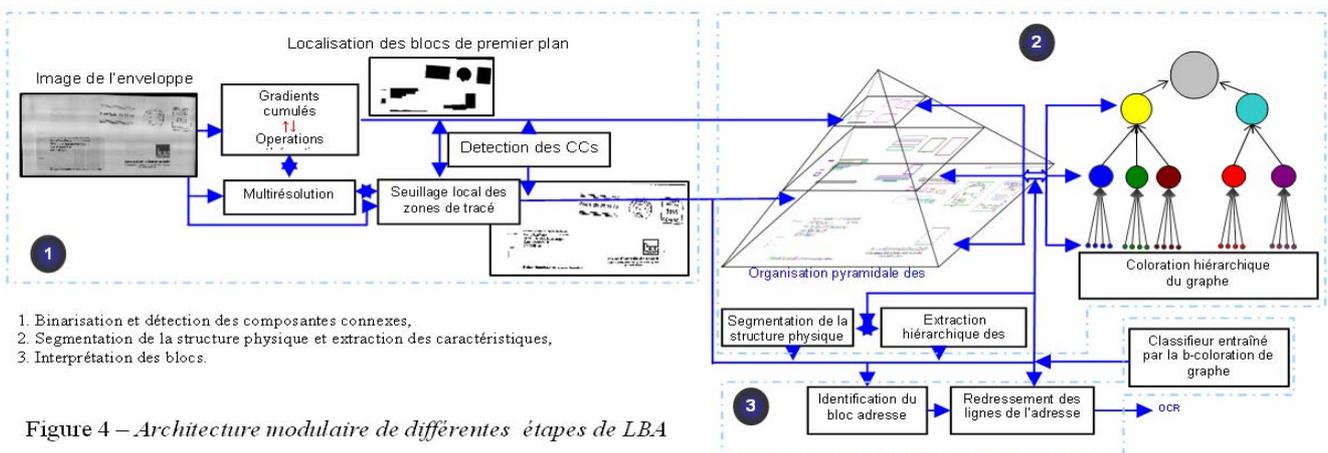


Figure 4 – Architecture modulaire de différentes étapes de LBA

sur une analyse hiérarchique à chaque niveau des rectangles englobants, lesquels constituaient une source importante d'information très souvent utilisée durant le processus de description.

### 3.2 Notre stratégie d'extraction des caractéristiques

Chaque bloc peut être décrit par un ensemble des caractéristiques issues de l'analyse hiérarchique des trois niveaux de la pyramide des données. La progression dans la hiérarchie permet, à chaque niveau, d'acquérir des connaissances plus précises sur le contenu de l'image.

A l'inverse de la plupart des méthodes existantes, notre idée étant de faire coopérer cette phase avec la phase de la segmentation physique tout en permettant, à un niveau de la hiérarchie, d'utiliser toutes les informations exprimées dans les autres niveaux (Figure 6).



Figure 6 – Extraction hiérarchique des caractéristiques

À l'aide de l'ACP nous avons pu sélectionner un nombre minimal de primitives tout en assurant une disparité maximale entre les objets de différentes natures. Si  $Vd_{L_k}(i)$  représente la description d'un objet  $i$  au niveau  $k$ , sa description totale est la combinaison des descriptions de trois niveaux  $L_1, L_2$ , et  $L_3$  (Formule 1).

$$Vd_{total}(i) = Vd_{L_1} \cup Vd_{L_2} \cup Vd_{L_3} \quad (1)$$

### 3.3 Segmentation de la structure physique basée sur la coloration des graphes

Les techniques de segmentation que nous avons étudiées ne peuvent pas, toujours, produire des blocs assez uniformes dans un environnement complexe. On parle de sur-segmentation lorsque l'élément constitutif est lui-même fragmenté, et de sous-segmentation lorsque plusieurs éléments constitutifs n'ont pas pu être isolés. Par conséquent, la présence des parasites ou une information incomplète dans un bloc adresse peut introduire des erreurs dans sa description et conduit à une mauvaise interprétation. Pour avoir plus de robustesse, nous avons introduit efficacement la CHG dans la phase de la segmentation. Notre idée est de former à chaque niveau de la hiérarchie des groupes homogènes de composantes pour aboutir à une description plus précise. Soit le graphe  $G_k(Vd_{L_k}, E_{L_k > S_k})$  défini par ses sommets

$$Vd_{L_k} = \{Vd_{L_k}(i) / i = 1 \dots N_k\} \quad (2)$$

Avec :

$$E_{L_k > S_k}(Vd_{L_k}(i), Vd_{L_k}(j)) = \begin{cases} 1 & \text{si } d_{i,j} > s_k \Leftrightarrow Vd_{L_k}(i) \text{ adjacent à } Vd_{L_k}(j) \\ 0 & \text{si non } \Leftrightarrow Vd_{L_k}(i) \text{ non adjacent à } Vd_{L_k}(j) \end{cases} \quad (3)$$

Tel que  $Vd_{L_k}$  représente les descripteurs de l'ensemble  $V(L_k)_{k=1,2 \text{ et } 3}$  de niveau  $k$  et les arcs  $E_{L_k > S_k}$  représentent les

dissemblances  $d_{i,j}$  entre chaque paire d'individus limité par le seuil  $S_k$ . Une méthodologie du calcul automatique des seuils est représentée en détails dans [15].

Au début, on colore le graphe  $G1(Vd_{L_1}, E_{>S1})$  par l'algorithme Coloration\_Graphe () donné ci-dessous. On superpose ainsi ces couleurs avec la carte des blocs  $V(L_3)$  pour fragmenter tout bloc qui correspond aux plusieurs couleurs et toute couleur qui correspond aux plusieurs blocs. En exploitant ces nouvelles connaissances, nous appliquons une coloration de graphe  $G3(Vd_{L_3}, E_{>S3})$  formé par  $Vd_{L_3}$  ensemble de descripteurs des fragments que nous fusionnons pour former une carte des blocs uniformes notée  $V^*(L_3)$ . Finalement, nous perfectionnons la description totale de chaque bloc par un nouveau jeu de caractéristiques extraites de la couche 2 issue d'une coloration du graphe  $G2(Vd_{L_2}, E_{>S2})$  avec  $Vd_{L_2}$  ensemble de descripteurs définis par l'analyse de  $V(L_1)$  et  $V^*(L_3)$ .

**Algorithme :** Coloration\_Graphe ()

Début

Si  $c(i) \neq \emptyset$  Alors

soit  $M := Nc(i) \cup \{c(i)\}; q := 0;$

Pour chaque sommet  $j \in N(i)$  sachant que  $c(j) := \emptyset$  faire

$q := \min \{k | k > q, k \notin M \text{ and } k \notin Nc(j)\};$

Si  $q \leq \Delta + 1$  Alors  $c(j) := q;$

Sinon  $c(j) := \min \{k | k \notin Nc(j)\};$  FinSi

Finpour

FinSi

Fin.

avec  $c(i)$  est la couleur de sommet  $i$ ,  $N(i)$  est l'ensemble de ses sommets adjacents ( $Nc(i)$  est l'ensemble des couleurs des sommets  $N(i)$ ),  $d(i) = |N(i)|$  son degré, et :

$$\Delta = \max \{d(i) | i \in V\} \quad (4)$$

### 3.4 Apprentissage basée sur la b-coloration

Afin de préparer une base d'apprentissage représentative nous avons sélectionné 400 blocs de plusieurs catégories (AB, timbre, logos...), issus de la segmentation physiques des images d'enveloppes à grande variété. Nous nous servons, de l'algorithme de b-coloration proposé dans [17] pour décomposer ces blocs en plusieurs classes. Ce procédé offre une bonne représentation des classes par les sommets dominants (représentants des classes) qui assurent une grande disparité interclasses (Figure7). Ces sommets seront utiliser, ainsi, pour LBA.

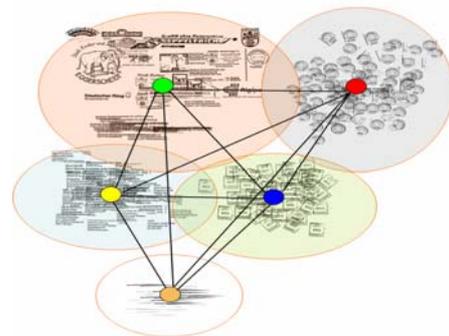


Figure 7 – Apprentissage : classification des blocs par la b-coloration et détections des représentants des classes

### 3.5 LBA par b-coloration des graphes

Pour sélectionner le BA dans une liste des blocs candidats  $S = \{V_{i=1..N}\}$  nous comparons dans l'espace des caractéristiques la description de chacun avec celles de tous les sommets dominants  $S^* = \{v_{j=1..M}\}$  issus de la phase d'apprentissage. L'algorithme d'appariement affecte à chaque bloc de  $S$  la même désignation de son plus proche sommet de  $S^*$ . De plus, cette interprétation fournit des nouvelles connaissances sur les relations spatiales entre les zones de l'enveloppe nécessaires à la prise d'une décision adéquate dans le cas de confusion. La dissemblance entre  $v_i$  et  $v_j^*$  est donnée par la distance généralisée de Minkowski d'ordre  $\alpha$  ( $\alpha \geq 1$ ).

$$d_{i,j} = \left( \sum_{k=1}^{N_f} g_k(v_i^k, v_j^{k*})^\alpha \right)^{\frac{1}{\alpha}} \quad (5)$$

Si  $\alpha = 1 \rightarrow d_{i,j}$  = la distance City Block, si  $\alpha = 2 \rightarrow d_{i,j}$  = distance Euclidienne et plus  $\alpha$  augmente, plus  $d_{i,j}$  tend vers la distance de Chebyshev.  $N_f$  est la longueur des vecteurs des caractéristiques.  $g_k$  est la fonction de dissemblance qui compare les caractéristiques deux à deux.

## 4 Expérimentation

Sur un corpus de 750 images d'enveloppes, le taux de LBA par notre méthode est de 98%. Afin de comparer les performances de notre nouveau système avec celles de l'ancien, nous avons cherché l'origine des rejets des 100 enveloppes au niveau de chaque phase de LBA. La courbe des rejets de l'ancien système montre que 18 enveloppes ont été rejetées à cause d'une mauvaise binarisation, 53 à cause d'une mauvaise segmentation physique et 29 à cause d'une mauvaise interprétation des blocs. Par la suite, nous avons appliqué notre méthode de LBA sur la même base: 7 enveloppes seulement ont été rejetées.

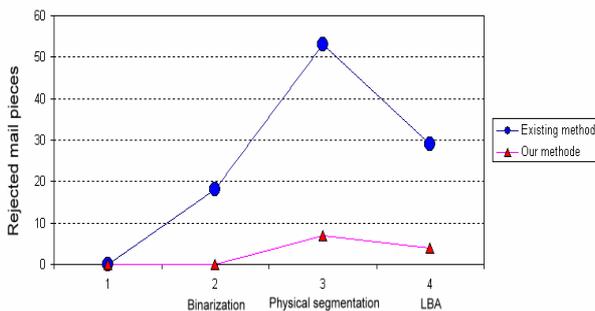


Figure 8 – Performances de notre méthode selon l'origine

## 5 Conclusion

Nous avons présenté dans cet article une nouvelle approche de localisation du bloc-adresse basée sur une coloration hiérarchique de graphe et une organisation pyramidale des données. Nous avons vu que la coloration hiérarchique de graphe a montrée une grande robustesse aux objets parasites considérés comme facteurs d'erreur de la segmentation physique, et que la b-coloration a conduit à une amélioration remarquable au niveau de

l'interprétation des blocs. De plus, nous avons pu augmenter les cohérences entre les différents modules et réduire les temps de calcul et les taux de rejet.

## Références

- [1] Wang Ching-Huei, P.W. Palumbo, S.N. Srihari, Object recognition in visually complex environments: an architecture for locating address blocks on mail pieces, Pattern Recognition, 1988., 9th International Conference, IEEE, 1988, vol.1, Pages: 365 – 367.
- [2] C Viard-Gaudin, D Barba, A multi-resolution approach to extract the address block on flat mail pieces, ICASSP-91, International Conference, vol.4, Pages: 2701 - 2704 .
- [3] B.Yu, A.K.Jain, M.Mohiuddin, Address block location on complex mail pieces, Document Analysis and Recognition, Fourth International Conference, IEEE, 1997, vol.2, Pages: 897 – 901.
- [4] Seon Hwa Jeong, Seung Ick Jang, Yun-Seok Nam, Locating destination address block in Korean mail images, ICPR 2004, IEEE, Vol.2, Pages: 387 – 390.
- [5] L.F.Eiterer, J.Facon, D.Menoti, Postal envelope address block location by fractal-based approach, Computer Graphics and Image Processing, 17th Brazilian Symposium, IEEE, 2004, Pages: 90 – 97.
- [6] N. Otsu, A threshold selection method from grey-level histogram, IEEE trans system, man and cybernetics, vol 9, pp. 62-66, 1979.
- [7] J. Sauvola, and al. Adaptive Document Binarization, ICDAR'97, vol 1, pp. 147–152, 1997.
- [8] Z. Pavlidis and J. Zhou, A Page Segmentation and Classification, CVGIP92, vol.54, no. 6, Pages: 484-496.
- [9] Regentova, E.; Latifi, S.; Shulan Deng; Dongsheng Yao, An Algorithm with Reduced Operations for Connected Components Detection in ITU-T Group 3/4 Coded Images. Pattern Analysis and Machine Intelligence, IEEE , Vol. 24, Pages:1039 – 1047.
- [10] Olivier Déforges and Dominique Barba, A fast multiresolution text-line and non text line structures extraction and discrimination scheme for document image analysis, ICPR 94, Pages: 134-138.
- [11] Shin-Ywan Wang; Yagasaki, T; Block selection: a method for segmenting a page image of various editing styles, ICDAR.1995, Vol 1, Pages:128 - 133 vol.1
- [12] Z. Shi; Venu Govindaraju, Line separation for complex document images using fuzzy runlength, Document Image Analysis for Libraries, DIAL 2004, Proceedings, First International Workshop on 2004 Pages:306 – 312
- [13] D. Drivas, A. Amin, Page segmentation and classification utilising a bottom-up approach, Document Analysis and Recognition, ICDAR.1995, Pages : 610 - 614 vol.2.
- [14] Brice Effantin and Hamamache Kheddouci , The b-chromatic number of power graphs, DMTCS 2003, Vol.6, pp. 45-54.
- [15] Haytham Elghazel, Mohand-Said Hacid, Hamamache Kheddouci and Alain Dussauchoy, A New Clustering Approach for Symbolic Data: Algorithms and Application to Healthcare Data, BDA 2006, Lille, France.
- [16] S. Corteel, M. Valencia-Pabon, J.C Vera, On approximating the b-chromatic number, Discrete Applied Mathematics archive, 2005, Vol146, Pages: 106 – 110.
- [17] Brice Effantin and Hamamache Kheddouci, a distributed algorithm for a b-coloring of a graph, ISPA'2006, Serrento, Italy, 2006 (IEEE).
- [18] Djamel Gaceb, Frank Lebourgeois, Véronique Eglin, Hubert Emptoz, Contribution to the automatic recognition of business documents, 6p, IWFHR 2006, La Baule, France.