

Séparation recto/verso d'un document par modélisation markovienne à double couche

Christian Wolf

LIRIS - Laboratoire d'informatique en images et systèmes d'information - UMR 5205

Université de Lyon

INSA-Lyon, Bât. J.Verne, 20, Av. Albert Einstein

69621 Villeurbanne cedex, France

christian.wolf@liris.cnrs.fr

Résumé

Nous proposons un modèle markovien à deux couches pour la séparation des deux faces d'un document, dont une seule face a été numérisé. A l'aide de deux champs de Markov séparés, un pour chaque face, chaque pixel est modélisé par deux variables cachées connectées par une unique variable observée. L'avantage de cette formulation est une meilleure adaptation au processus ayant créé l'image observée (la superposition de deux pages indépendantes) ainsi que l'amélioration de la restauration, ç.à.d. de l'estimation des pixels recto, par une estimation des pixels verso couverts par ce derniers. L'inférence des variables cachées est réalisée par un algorithme itératif à base de coupure minimale dans un graphe étendant l'algorithme d'expansion alpha. Les résultats sont évalués à la fois de façon empirique ainsi que par l'amélioration d'un résultat de reconnaissance OCR.

Mots clefs

Restauration d'images de documents, MRF, graph cuts

1 Introduction

Dans ce travail, nous traitons le problème de la suppression de la face verso d'un document visible dans une image après la numérisation de la face recto. Nous supposons que la numérisation de la face verso n'est pas disponible (séparation aveugle). Dans ce cas, la tâche peut se résumer à un problème de segmentation : classement de chaque pixel en tant que *recto*, *verso*, *fond*, et éventuellement *recto-et-verso* (en même temps).

Plusieurs techniques ont été proposés dans la littérature pour ce type de problèmes, par exemple l'analyse de composantes indépendantes [1, 2], les ondelettes [3] et d'autres approches fréquentielles [4], les équations à dérivées partielles [5], des approches non-aveugles nécessitant l'alignement des pages [6, 7, 8] et le seuillage guidé par modèle [9].

La régularisation par champs de Markov (MRF) a déjà été appliquée à ce problème [10, 11, 12]. Elle permet de créer un modèle statistique exploitant des connaissances

sur le processus de dégradation ainsi que sur le contenu de l'image (la connaissance *a priori*). Toutefois, les méthodes précédentes traitent la séparation recto/verso de la même manière que la segmentation classique d'images. Dans ce travail, nous montrons que les performances peuvent être améliorées de façon significative lorsque la connaissance de l'existence de deux différentes parties du document est prise en compte et modélisée en conséquence.

Ce document est organisé comme suit. La section 2 introduit le modèle graphique, ç.à.d. le graphe de dépendances de l'ensemble des variables cachées — recto et verso — ainsi que des variables observées. La section 3 décrit la procédure d'estimation de paramètres et la section 4 donne une esquisse de la méthode d'inférence par coupure minimale dans un graphe. Enfin, la section 5 donne les résultats expérimentaux et la section 6 conclut.

2 Le modèle graphique

Les approches de segmentation par MRF sont souvent motivées par des hypothèses d'homogénéité, ç.à.d. des zones homogènes sont considérées comme étant plus probables que des changements fréquents d'étiquettes entre pixels voisins. Ce n'est pas justifié à tous les endroits de l'image observée lorsque celle-ci est le résultat d'une superposition de deux ou plusieurs "sources". Dans ce cas, les sources étant indépendantes l'une de l'autre, la connaissance *a priori* peut être disponible pour chaque source de l'image, mais pas pour le mélange.

Pour cette raison nous proposons un modèle *a priori* ayant deux champs d'étiquettes : un pour le côté recto (F^1) et un pour le côté verso (F^2), résultant en un modèle avec deux variables cachées pour chaque pixel (*recto* et *verso*) et un espace de configuration de deux valeurs pour chaque variable cachée (*texte* et *fond*). Les avantages de cette formulation sont les suivants :

- la connaissance *a priori* est appliquée à la partie du modèle pour laquelle elle a été obtenu, ç.à.d. pour les sources du mélange. Les hautes fréquences introduites par le mélange ne sont pas lissées par la régularisation markovienne.

– Une estimation des pixels verso recouverts par des pixels recto de type “texte”, possible uniquement avec deux champs d’étiquettes séparés, n’est pas seulement souhaitable dans le cas où une estimation de la page verso est demandée par l’application. Plus encore, une estimation correcte des pixels verso, par l’intermédiaire des interactions spatiales encodées par les deux champs d’étiquettes, aide à estimer correctement le pixel verso non couverts par un pixel recto de type “texte”. Cela améliore les performances de l’algorithme en améliorant l’estimation de la face recto.

Un MRF [13, 14] est un modèle causal sur un graphe non orienté considérant une image comme un processus stochastique et assignant une probabilité à chaque réalisation d’un ensemble de variables f reliées aux sommets du graphe. La distribution de probabilité jointe des variables suit une distribution de Gibbs [15] définie sur les cliques maximales du graphe :

$$P(f) = \frac{1}{Z} \exp \{-U(f)/T\} \quad (1)$$

où Z est une constante de normalisation, T est un facteur de température (très souvent = 1), $U(f) = \sum_{c \in \mathcal{C}} V_c(f)$ est une fonction d’énergie définie selon l’application, \mathcal{C} est l’ensemble des cliques du graphe et $V_c(f)$ est la fonction d’énergie pour une réalisation du champs f et définie sur une seule clique c .

Très souvent, un MRF modélise la connaissance *a priori* sur des variables cachées, nécessitant de le combiner, dans le cadre de l’estimation Bayésienne, avec un terme d’attache aux données impliquant les variables cachées ainsi que les variables observées. Etant donnée la nature de notre problème, nous préférons interpréter l’ensemble complet des variables cachées et observées comme un seul MRF. Dans la suite de cet article nous considérons donc un graphe $\mathcal{G} = \{V, E\}$ avec un ensemble de sommets V et un ensemble d’arêtes E , où V est partitionné en trois sous-ensembles distincts : les deux champs cachés F^1 et F^2 et un champs observé D correspondant à l’image d’entrée. Les trois champs sont indexés par les mêmes indices correspondant aux pixels de l’image, ç.à.d. F_s^1 , F_s^2 et D_s désignent, respectivement, l’étiquette cachée recto, l’étiquette cachée verso et l’observation du même pixel s .

Le graphe de dépendances proposé (voir figure 1) contient les types de cliques suivants : les cliques d’ordre un et deux du sous graphe F^1 ainsi que les mêmes types de cliques du sous graphe F^2 (nous supposons que les potentiels d’énergie des cliques à trois sommets sont égaux à zéro) et enfin les cliques “inter-champs” entre F^1 , F^2 et D contenant trois sommets, un de chaque champs.

La distribution jointe du graphe peut donc être présentée

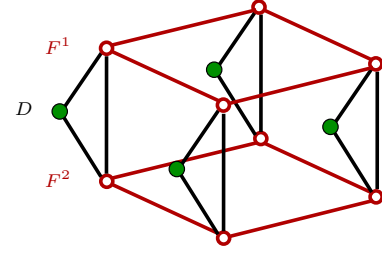


Figure 1 – Le graphe de dépendances du modèle créé pour une image de 4 pixels.

comme suit :

$$\begin{aligned} P(f^1, f^2, d) &= \frac{1}{Z} \exp \{-U(f^1) - U(f^2) - U(f^1, f^2, d)\} \\ &= \frac{1}{Z} \exp \{-U(f^1) - U(f^2)\} \exp \{-U(f^1, f^2, d)\} \\ &= P(f^1, f^2) P(d|f^1, f^2) \end{aligned} \quad (2)$$

La dernière égalité donne une indication de l’interprétation Bayésienne du problème : le premier facteur correspond à la connaissance *a priori* et le deuxième facteur correspond à la vraisemblance des données déterminée par le modèle d’observation et de dégradation. Comme on peut constater, la probabilité *a priori* est le produit de deux probabilités associées aux deux champs F^1 and F^2 :

$$P(f^1, f^2) = P(f^1)P(f^2) \quad (3)$$

En d’autres termes, l’écriture sur le recto est indépendante de l’écriture sur la page verso, ce qui prend tout son sens puisque les deux pages ne sont pas nécessairement influencées l’une par l’autre — elles peuvent même avoir été créés par des auteurs différents. Toutefois, cette indépendance ne concerne que la situation dans laquelle aucune observation n’a été faite. En présence d’observations (l’image numérisée), les deux champs cachés ne sont pas indépendants en raison des cliques de trois sommets impliquant des paires de variables cachées et une variable observée. Intuitivement parlant, cela peut être illustré par l’exemple suivant : si l’observation d’un pixel donné suggère qu’au moins un des deux côtés du document contient du texte à cet endroit (si, par exemple, le niveau de gris du pixel est plutôt faible pour un document à texte noir sur fond blanc), alors la connaissance que l’étiquette recto est égale à *fond* va augmenter la probabilité que le pixel verso sera *texte*.

Les termes $U(f^1)$ and $U(f^2)$ correspondent à deux mo-

dèles de type Potts, un pour chaque champs :

$$\begin{aligned}
U(f^1) + U(f^2) &= U(f^1, f^2) \\
&= \sum_{\{s\} \in \mathcal{C}_1} \alpha^1 f_s^1 + \sum_{\{s, s'\} \in \mathcal{C}_2} \beta_{s, s'}^1 \delta_{f_s^1, f_{s'}^1} \\
&+ \sum_{\{s\} \in \mathcal{C}_1} \alpha^2 f_s^2 + \sum_{\{s, s'\} \in \mathcal{C}_2} \beta_{s, s'}^2 \delta_{f_s^2, f_{s'}^2}
\end{aligned} \quad (4)$$

où \mathcal{C}_1 est l'ensemble de cliques à un sommet unique, \mathcal{C}_2 est l'ensemble de cliques de taille 2 (paires) et δ est le symbole de Kronecker défini comme $\delta_{i, j} = 1$ si $i = j$ et 0 sinon. Les paramètres $\alpha^1, \alpha^2, \beta^1$ et β^2 servent à gérer la connaissance *a priori* sur l'équilibre entre les étiquettes *texte* et *fond* ainsi que sur les forces d'interaction entre les pixels voisins. Le terme $U(f^1, f^2, d)$ est relié à la vraisemblance des données :

$$\begin{aligned}
U(f^1, f^2, d) &= -\log p(d|f^1, f^2) \\
&= \sum_s -\log \mathcal{N}(d_s; \mu_s, \Sigma_s)
\end{aligned} \quad (5)$$

où μ_s est la moyenne de la classe f_s et Σ_s est la matrice de covariance pour la classe f_s donnée comme ci-dessous :

$$\begin{aligned}
\mu_s &= \begin{cases} \mu_r & \text{if } f_s^1 = \textit{text} \\ \mu_v & \text{if } f_s^1 = \textit{background} \text{ and } f_s^2 = \textit{text} \\ \mu_{bg} & \text{else} \end{cases} \\
\Sigma_s &= \begin{cases} \Sigma_r & \text{if } f_s^1 = \textit{text} \\ \Sigma_v & \text{if } f_s^1 = \textit{background} \text{ and } f_s^2 = \textit{text} \\ \Sigma_{bg} & \text{else} \end{cases}
\end{aligned} \quad (6)$$

où μ_r, μ_v et μ_{bg} sont, respectivement, et *dans l'image dégradée*, la moyenne de la classe recto, de la classe verso, et de la classe fond. Les matrices de covariance sont définies de façon similaire.

3 Estimation des paramètres

Dans ce travail, nous avons choisi d'estimer les paramètres d'une manière supervisée sur une image segmentée initialement avec la méthode k-means ($k = 3$) puis post-traitée avec un filtre médian. Les étiquettes obtenues par la méthode k-means étant non ordonnées et non pourvues de signification, la reconnaissance de leur type (*recto*, *verso*, *fond*) s'avère nécessaire. Nous avons conçu un algorithme capable de détecter ce type sans nécessité de recourir aux informations provenant de l'histogramme de l'image, ç.à.d. sans connaissance *a priori* sur la "couleur" (le niveau de gris) du texte. Notre méthode se sert de statistiques sur les changements d'étiquettes entre pixels voisins ainsi que d'une analyse de composantes connexes. En supposant que les composantes connexes recto ont tendance à couvrir, et donc de séparer en plusieurs parties, d'autres éléments, la signification des étiquettes peut être déterminée.

Les paramètres du modèle de Potts ($\alpha^1, \alpha^2, \beta^1, \beta^2$) sont estimés avec une méthode par moindres carrés proposée par Derin et al. [16]. Les champs recto étant plus stable, nous estimons les paramètres sur ce champs uniquement, car toutes ses étiquettes sont directement liées à l'image observée. Les paramètres du champs verso sont calculés à partir des paramètres du champs recto en supposant que, statistiquement parlant, le champs verso est une version inversée du champs recto. Les paramètres du modèle d'observation sont estimés à l'aide des estimateurs classiques de type vraisemblance maximale (la moyenne empirique et les covariances empiriques).

4 Inférence

Pour la minimisation de l'énergie postérieure nous avons mené des expériences avec deux méthodes différentes : le recuit simulé [13] et une adaptation de la méthode par coupe minimale dans un graphe [17] appliqué au graphe de dépendance de notre modèle. Nous avons étendu l'algorithme *α -expansion move* proposé par Boykov et al. [18] et amélioré par Kolmogorov et al. [17]. Cet algorithme est capable de trouver, de façon rapide, une bonne solution pour certains problèmes d'étiquetage, sous condition que la fonction d'énergie à minimiser satisfait certaines propriétés. Plus spécifiquement, pour les termes binaires, ç.à.d. les interactions entre deux variables $U(\cdot, \cdot)$, la condition suivante doit être satisfaite [17] :

$$U(0, 0) + U(1, 1) \leq U(0, 1) + U(1, 0) \quad (7)$$

Comme on peut constater, cette condition, désignée "régulière" [17], est naturellement satisfaite pour les potentiels d'énergie des cliques du modèle de Potts. Par contre, les potentiels d'énergie log-Gaussiens issus de la vraisemblance des données (voir équation (5)) ne sont pas réguliers. En effet, la régularité de ces fonctions à trois variables, dont les deux variables cachées f_s^1 et f_s^2 , dépend de l'observation d_s sur chaque site. L'énergie de ce type de clique va donc être régulière pour certains pixels et non régulière pour d'autres.

Pour palier à ce problème nous avons conçu un algorithme itératif alternant entre une estimation des étiquettes cachées du champs recto f^1 en figeant les étiquettes f^2 du champs verso et vice versa. Les étiquettes figées étant constantes, les fonctions d'énergie correspondantes sont désormais unaires, ç.à.d. dépendant d'une seule variable cachée, et donc naturellement régulières. Toutefois nous pouvons améliorer l'algorithme en figeant seulement les variables cachées pour les pixels dont les termes de vraisemblance le demandent, ç.à.d. les pixels dont la fonction d'énergie n'est pas régulière. Pour les autres, les deux variables cachées, recto et verso, peuvent être estimées en même temps et de façon optimale.

Une esquisse de l'algorithme est donnée dans la figure 2, plus de détails sur sa dérivation sont donnés dans [19].

Entrées : d (une réalisation du champs observé)
Sorties : f^1, f^2 (les champs d'étiquettes estimés)
 $f^1, f^2 \leftarrow$ Initialiser les champs d'étiquettes (e.g. avec la méthode k-means)
 $H \leftarrow$ déterminer les sommets s réguliers ($H_s = 1 \Leftrightarrow s$ est régulier)
répéter
- Fixer les f_s^1 pour tous les s et les f_s^2 pour les $s : H_s = 0$, puis estimer de façon optimale les f^2 pour tous les s et les f_s^2 pour $H_s = 1$
- Fixer les f_s^2 pour tous les s et les f_s^1 pour les $s : H_s = 0$, puis estimer de façon optimale les f^1 pour tous les s et les f_s^1 pour $H_s = 1$
jusqu'à convergence

Figure 2 – L'algorithme d'inférence optimisant, de façon itérative et optimale, deux sous problèmes binaires. H est une matrice stockant pour chaque pixel s si son terme de vraisemblance est régulier ($H_s = 1$) ou non ($H_s = 0$).

Niveau de bruit	$\sigma=10$	$\sigma=15$	$\sigma=20$
K-Means	0.25	1.40	3.56
MRF simple	0.03	0.23	0.73
MRF double	0.01	0.08	0.31

Tableau 1 – Erreur de classification (en % de pixels) sur des images synthétiques avec des quantités de bruit différentes.

5 Expériences

Pour évaluer notre algorithme, nous avons effectué des expériences à la fois sur des données synthétiques ainsi que sur des données réelles. Dans tous les cas, les images ont été converties en niveaux de gris avant le traitement.

Images synthétiques Nous avons créé des images de synthèse, conformément au modèle de dégradation décrit dans la section 2. Pour chaque image de synthèse, deux images parfaites ont été superposées et du bruit Gaussien avec des variances différentes a été ajouté. La table 1 montre les résultats de notre algorithme ainsi que d'autres algorithmes : une simple segmentation par k-means, une segmentation de type MRF simple, ç.à.d. avec un seul champs d'étiquettes, et l'algorithme proposé à double champs d'étiquettes. Nous voyons que le MRF double surpasse les autres méthodes.

Documents numérisées - restauration La méthode a également été testée sur des images réelles, imprimées ou manuscrites, dont des exemples sont montrés dans la figure 3. Comme on peut le constater, la régularisation MRF est capable de supprimer de nombreux artefacts présents dans l'image ségmentée par la méthode k-means. La méthode de double MRF diminue encore plus le nombre d'artefacts.

Méthode	1 ^{ier}	2 ^e	3 ^e
K-Means	18	10	36
Single-MRF	13	39	12
Double-MRF	33	15	16
Total	64	64	64

(a)

Méthode	1 ^{ier}	2 ^e
K-Means	21	43
Double-MRF	43	21
Total	64	64

(b)

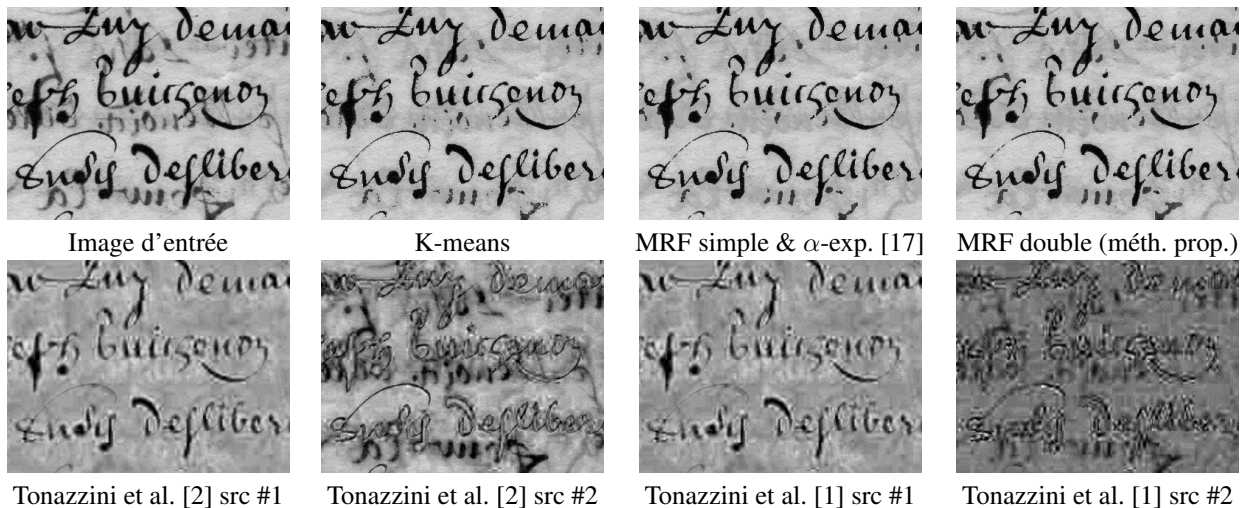
Tableau 2 – Résultats du test empirique obtenus avec inférence par recuit simulé : (a) toutes les méthodes (b) en ignorant la méthode classée troisième (MRF simple).

La vérité terrain étant indisponible pour ce type d'images, nous avons présenté un ensemble de 4 images d'entrée et de 3 images restaurées pour chaque image d'entrée à 16 personnes différentes (après avoir changé l'ordre des images de façon aléatoire, bien évidemment) et nous leur avons demandé de faire un classement par qualité perçue. Les résultats de ces $N = 64$ tests sont indiqués dans le tableau 2a. La méthode proposée a été classé au premier rang 33 fois contre 18 fois (k-means) et 13 (MRF simple).

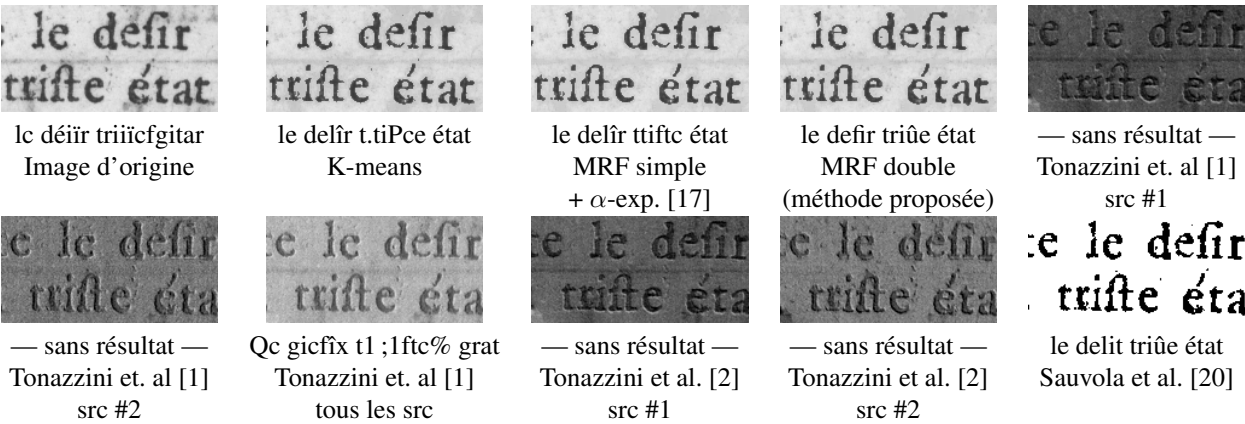
La performance relativement surprenante de l'algorithme k-means est dû au fait qu'il tend à conserver une plus grande quantité de pixels recto que le MRF simple. Pour tester la signification statistique de ce résultat, nous supposons l'hypothèse nulle suivante : $H_0 =$ "la méthode MRF double est aussi efficace que les deux autres méthodes".

Nous pouvons déjà conclure à partir des données que la méthode n'est pas moins efficace, il suffit donc de rejeter H_0 . Notre statistique de test sera $U =$ "le nombre de fois où la méthode MRF double est classée au premier rang". En supposant H_0 vrai, la probabilité d'une méthode d'être classée au premier rang pour une image est de $\pi = \frac{1}{3}$, la statistique U suit donc une distribution binomiale $\mathcal{B}(64, \frac{1}{3})$. La probabilité d'obtenir la valeur $U=33$ ou une valeur plus extrême est de 0,00197. Supposons un test unilatéral avec un niveau de signification de $\alpha=0,025$, H_0 doit être rejetée.

Afin de prouver que la méthode proposée est meilleure que la méthode classée deuxième, nous considérons les résultats sans tenir compte de la méthode classée au troisième rang, ce qui nous donne les résultats indiqués dans le tableau 2b. La nouvelle hypothèse nulle sera la suivante : $H_0 =$ "La méthode MRF double est aussi efficace que la méthode k-means". La statistique de test U suit une distribution binomiale $\mathcal{B}(64, 0.5)$, la probabilité $P(U \geq 43)$ est égale à 0,004073, ce qui nous amène à rejeter H_0 . Le MRF double est donc plus efficace que la méthode k-Means.



(a)



(b)

Figure 3 – Résultats de restauration pour des documents manuscrits (a) et imprimés (b), avec des résultats OCR sur les documents imprimés. Tous les résultats de type MRF ont été obtenu avec optimisation par coupure minimale.

Documents numérisés - OCR Le logiciel OCR Google Tesseract a été appliqué à un ensemble de 104 pages de faible qualité soumis à plusieurs algorithmes de restauration concurrents (voir la figure 3b). La méthode proposée donne les meilleurs résultats avec un rappel de 83,23 % et une précision de 74,85 % (au niveau des caractères) contre le concurrent le plus proche, le MRF simple, parvenant à 81,99 % et 72,12 %. Plus de détails sur les expériences sont donnés dans [19].

6 Conclusion

Dans cet article, nous avons présenté une méthode permettant de séparer le côté recto d'un document du côté verso en se servant d'une seule image numérisée. La nouveauté de la méthode est la séparation du MRF en deux champs d'étiquettes différents, chacun régularisant une des deux faces du document. Cette séparation permet d'estimer les pixels verso du document couverts par les pixels recto de type "texte", ce qui améliore de façon significative la performance de l'algorithme. La méthode a été validé empiriquement,

ainsi que par l'amélioration de résultats de reconnaissance.

7 Remerciements

Nous remercions Prof. Anna Tonazzini de nous avoir fourni les codes sources des deux méthodes basées sur la séparation de sources, de son aide précieuse pour la mise en place des expériences ainsi que pour les discussions très intéressantes.

Références

- [1] A. Tonazzini et L. Bedini. Independent component analysis for document restoration. *I.J. on Doc. Anal. and Rec.*, 7(1) :17–27, 2004.
- [2] A. Tonazzini, E. Salerno, et L. Bedini. Fast correction of bleed-through distortion in grayscale documents by a blind source separation technique. *I.J. on Doc. Anal. and Rec.*, 10(1) :17–25, 2007.
- [3] Q. Wang, T. Xia, C.L. Tan, et L. Li. Directional wavelet approach to remove document image interference.

Type de méthode	Méthode		Rappel (en %)	Précision (en %)	Coût (abs.)	Nombre d'images (en %)
—	Sans restauration		65.65	49.91	76,752	100
Sans contexte	Niblack [21] (segm. uniquem.)		-	-	-	-
	Sauvola et al. [20] (segm. uniquem.)		78.75	66.78	45,363	100
	K-Means (k=3)		79.82	68.43	42,675	100
Sép. de sources	Tonazzini et al. [2] - src #1	‡	41.00	30.05	74,819	66
	Tonazzini et al. [2] - src #2	†	-	-	-	-
	Tonazzini et al. [1] - src #1	†	-	-	-	-
	Tonazzini et al. [1] - src #2	†	-	-	-	-
	Tonazzini et al. [1] - 3 sources	‡	50.52	33.90	101,280	89
MRF	MRF simple & α -exp. move [17]		81.99	72.12	36,744	100
	MRF double (méthode proposée)		83.23	74.85	32,537	100

† indisponible : performance du OCR non suffisant pour une évaluation correcte
‡ résultats obtenus avec un sous ensemble d'images (coût abs. non comparable).

Tableau 3 – Résultats de reconnaissance OCR sur une base de 104 images de documents numérisées : sans restauration et après restauration avec des méthodes différentes.

- Dans *I.C. on Doc. Anal. and Rec.*, pages 736–740, 2003.
- [4] H. Nishida et T. Suzuki. Correcting show-through effects on document images by multiscale analysis. Dans *Proc. of the I.C. on Pattern Recognition*, volume 3, pages 65–68, 2002.
- [5] F. Drira, F. Lebourgeois, et H. Emptoz. OCR accuracy improvement through a PDE-based approach. Dans *I.C. on Doc. Anal. and Rec.*, volume 2, pages 1068–1072, 2007.
- [6] G. Sharma. Show-through cancellation in scans of duplex printed documents. *IEEE Tr. on Image Processing*, 10(5) :736–754, 2001.
- [7] E. Dubois et A. Pathak. Reduction of bleed-through in scanned manuscript documents. Dans *Proc. of the Image Processing, Image Quality, Image Capture Systems Conf.*, pages 177–180, 2001.
- [8] C.L. Tan, R. Cao, et P. Shen. Restoration of archival documents using a wavelet technique. *IEEE Tr. on PAMI*, 24(10) :1399–1404, 2002.
- [9] H.-S. Don. A noise attribute thresholding method for document image binarization. *I.J. on Doc. Anal. and Rec.*, 4(2) :131–138, 2000.
- [10] A. Tonazzini, S. Vezzosi, et L. Bedini. Analysis and recognition of highly degraded printed characters. *I.J. on Doc. Anal. and Rec.*, 6(4) :236–247, 2003.
- [11] K. Donaldson et G.K. Myers. Bayesian super-resolution of text in video with a text-specific bimodal prior. *I.J. on Doc. Anal. and Rec.*, 7(2-3) :159–167, 2005.
- [12] C. Wolf et D. Doermann. Binarization of Low Quality Text using a Markov Random Field Model. Dans *Proc. of the I.C. on Pattern Recognition*, volume 3, pages 160–163, 2002.
- [13] S. Geman et D. Geman. Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images. *IEEE Tr. on PAMI*, 6(6) :721–741, 11 1984.
- [14] S.Z. Li. *Markov Random Field Modeling in Image Analysis*. Springer Verlag, 2001.
- [15] J.M. Hammersley et P. Clifford. Markov fields on finite graphs and lattices. unpublished manuscript, 1968.
- [16] H. Derin et H. Elliott. Modeling and Segmentation of Noisy and Textured Images Using Gibbs Random Fields. *IEEE Tr. on PAMI*, 9(1) :39–55, 1987.
- [17] V. Kolmogorov et R. Zabih. What energy functions can be minimized via graph cuts ? *IEEE Tr. on PAMI*, 26(2) :147–159, 2004.
- [18] Y. Boykov et V. Kolmogorov. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *Pattern Analysis and Machine Intelligence, IEEE Tr. on*, 26(9) :1124–1137, 2004.
- [19] C. Wolf. An iterative graph cut optimization algorithm for a double mrf prior. Rapport technique LIRIS RR-2008-017, Laboratoire d'Informatique en Images et Systèmes d'Information, INSA de Lyon, France, 2008.
- [20] J. Sauvola, T. Seppänen, S. Haapakoski, et M. Pietikäinen. Adaptive Document Binarization. Dans *International Conference on Document Analysis and Recognition*, volume 1, pages 147–152, 1997.
- [21] W. Niblack. *An Introduction to Digital Image Processing*, pages 115–116. Englewood Cliffs, N.J. : Prentice Hall, 1986.