

Extraction de caractéristiques dans les images par transformée multi-échelle

Guillaume JOUTEL, Véronique ÉGLIN, Stéphane BRÈS, Hubert EMPTOZ

LIRIS, INSA-Lyon, 20 avenue Albert Einstein, 69621 Villeurbanne Cedex

{guillaume.joutel, veronique.eglin, stephane.bres, hubert.emptoz}@liris.cnrs.fr

Résumé – Dans cet article nous présentons une nouvelle méthode d'extraction de caractéristiques d'images basée sur une transformée multi-échelle, la transformée en Curvelets. Cette dernière possède à la fois les propriétés de caractériser les textures des images selon différents niveaux d'échelle et de rendre compte de la grande variété des formes et de leur forte anisotropie même dans des environnements bruités (fond bruité, formes incomplètes). Les caractéristiques extraites de cette représentation sont la courbure et l'orientation sur le contour des formes, ces deux dimensions étant primordiales pour nos travaux. Les résultats obtenus sont présentés dans le cadre d'une application à la recherche d'informations sur des manuscrits anciens numérisés.

Abstract – The aim of this work is to extract features of an image from the result of a multi-scale transform, the Curvelets transform. This one is able in the same time to characterize textures of images according to various scale levels and to take into account the great diversity of shapes and their strong anisotropy even in a disturbed environments (disturbed background, incomplete shapes). Extracted features are curvature and orientation of shapes' contours. Those features are essential in our work. Results are presented in the information retrieval context over ancient manuscripts.

1 Introduction

Depuis leur introduction au début des années 1980, les ondelettes ont fait l'objet de beaucoup d'attention dans des domaines aussi diversifiés que le débruitage, la compression, le codage, l'imagerie médicale ou satellitaire . . .

Elles y ont démontré leur force mais les ondelettes séparables sont isotropes et ne peuvent donc pas capturer, par exemple, la régularité d'un contour dans une image. Ceci est dû au fait que les ondelettes sont des outils adaptés à la description des discontinuités de signaux mono dimensionnels et que cette propriété n'est plus vraie pour des dimensions supérieures. En particulier, l'orthogonalité de la décomposition et l'échantillonnage critique font apparaître des effets d'aliasing visibles autour des contours. De plus, le nombre d'orientations est limité et fixe, et les contours sont redondants d'un niveau de résolution à un autre, ce qui requiert un grand nombre de coefficients d'ondelettes pour les représenter.

Afin de pallier à ce problème, de nouvelles décompositions multi-résolution mieux adaptées à la représentation de tels signaux ont été introduites. Nous présentons dans la partie 2 certaines de ces représentations en nous attachant plus particulièrement dans la partie 3.1 à la description de l'une d'entre elle qu'est la transformée en Curvelets.

Nous verrons ensuite comment à partir de cette transformée en Curvelets nous avons pu extraire deux dimensions caractéristiques que sont la courbure et l'orientation le long des contours d'une forme. Ces deux dimensions étant primordiales dans le cadre de nos travaux portant sur l'analyse de documents manuscrits anciens mais pouvant tout autant être calculées sur des images naturelles.

2 État de l'art

Les transformées en Ondelettes et les autres décompositions multi-résolutions classiques (pyramides laplaciennes par exemple) semblent former une catégorie restreinte et limitée des possibilités de représentations multi-échelles de signaux multidimensionnels. Des travaux récents montrent qu'il est possible de définir des cadres théoriques de représentations multi-échelles plus vastes donnant naissance à de nouvelles transformées plus intéressantes et plus adaptées à l'extraction de structures géométriques lisses et continues telles que les contours d'objets [5]. C'est le cas de la nouvelle famille des transformées en Ridgelets, en Contourlets et en Curvelets que nous allons rapidement présenter ici. Ce sont des décompositions multi-échelles qui opèrent selon une multitude d'orientations fréquentielles issues d'une décomposition non standard de l'espace des fréquences des images et qui offrent un bon compromis entre la représentation éparse de traits caractéristiques et la qualité des images reconstruites.

Parmi ces nouvelles transformées multi-résolution cherchant à pallier aux problèmes induits par la transformée en ondelettes on peut abusivement distinguer deux classes : les transformées géométriques et les autres représentations qui s'attachent plus à l'amélioration d'un « défaut » particulier des ondelettes plus qu'à une modification conceptuelle.

2.1 Les transformées géométriques

Dans la classe des transformées géométriques on trouve deux grandes approches : l'approche non-adaptative et l'approche adaptative. Les transformées géométriques non adaptatives utilisent des bases de fonctions fixes mais elles

vont chercher à corriger l'aspect isotrope des transformées en ondelettes « classiques ». On trouve dans cette famille la transformée en Contourlets introduite par Minh Do et Martin Vetterli [2]. Elle est construite en combinant successivement deux étages de décomposition distincts : une décomposition multi-échelle suivie d'une décomposition directionnelle. On trouve également les travaux de Emmanuel Candès et Dave Donoho sur la transformée en Ridgelets [6] dans un premier temps et sur la transformée en Curvelets [4] dans un second temps. Ces deux transformées seront présentées dans la partie 3.1.

Dans le cas des transformées géométriques adaptatives, on cherche à définir une base dont les fonctions sont choisies au mieux pour s'adapter à une image donnée. Pour cela, une étape préalable d'estimation de la géométrie de l'image est nécessaire avant de procéder à la décomposition. Dans cette famille on trouve par exemple la transformée en Bandelettes [3] dans laquelle les bases sont obtenues par déformation d'une base d'ondelettes selon le flux géométrique local¹.

2.2 Autre approche d'amélioration des ondelettes classiques

L'une des faiblesses de la transformée en ondelettes classique est qu'elle n'est pas invariante par translation. Cela peut poser problème dans le cas par exemple de la détection de mouvements dans le domaine transformé. Pour pallier à cette faiblesse, Kingsbury [1] propose une représentation en arbre dual qui fournit une transformée redondante approximativement invariante par translation. Pour cela, Kingsbury propose d'appliquer d'une part une décomposition « classique » d'un signal en ondelettes et d'autre part, une décomposition de ce même signal avec les mêmes filtres mais déphasés d'un échantillon. Dans le cas mono-dimensionnel, cette représentation est redondante d'un facteur 2 et peut être assimilée à une transformation en ondelettes complexes où la partie réelle est fournie par la décomposition en ondelettes et la partie imaginaire est donnée par les coefficients issus de la décomposition déphasée.

3 Extraction des caractéristiques

Dans le cadre de nos travaux sur les manuscrits anciens, nous avons choisi de travailler avec la transformée en Curvelets. Avant d'en arriver à la méthode d'extraction de caractéristiques que nous avons développée, nous allons présenter une vision possible de la transformée en Curvelets avec dans un premier temps la transformée en Ridgelets.

3.1 De la transformée en Ridgelets à la transformée en Curvelets

La transformée en Ridgelets a été élaborée dans le but de capturer la régularité présente le long des contours rec-

tilignes. Pour ce faire, on utilise une transformée de Radon [8] qui permet de représenter une image de façon bijective dans le domaine polaire. On obtient ainsi une conversion des singularités rectilignes dans le domaine spatial en singularités ponctuelles dans le domaine de Radon. Cela peut se traduire pour une image $f \in L^2(\mathbb{R}^2)$ en :

$$R_f(\theta, t) = \int_{\mathbb{R}^2} f(x, y) \delta(x \cos(\theta) + y \sin(\theta) - t) dx dy \quad (1)$$

où $R_f(\theta, t)$ représente la projection radiale de f sur la droite d'équation $x \cos(\theta) + y \sin(\theta) = t$. La transformée en Ridgelets s'obtient alors en appliquant une transformée en ondelettes 1D le long de $R_f(\theta, \cdot)$ en utilisant la variable d'intégration t . La décomposition de f s'écrit alors :

$$RT_f(a, b, \theta) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} \psi\left(\frac{t-b}{a}\right) R_f(\theta, t) dt \quad (2)$$

où a est un facteur d'échelle, b un paramètre de translation, θ l'angle de projection et ψ une ondelette.

Malheureusement, avec une telle construction, la transformée en Ridgelets ne s'avère efficace que pour la caractérisation des contours rectilignes. On peut toutefois considérer qu'une image comporte, de manière locale, des contours rectilignes et c'est ce qui nous amène à la transformée en Curvelets. Cette transformée est obtenue en appliquant une transformée en Ridgelets dans une fenêtre de Fourier à support fini avec une dilatation de la fonction d'onde de a/a^2 . Une fois la transformée en Curvelets appliquée à notre image, il est possible de calculer la courbure et l'orientation sur les contours des formes.

3.2 Calcul de l'orientation et de la courbure

Grâce à la transformée en Curvelets que nous venons de décrire, nous obtenons pour chaque pixel des contours d'un objet de l'image un nombre de coefficients caractéristiques dépendant du niveau d'échelle et du nombre d'angles associés. Pour chacun de ces coefficients on évalue leur représentativité en comparaison avec les autres coefficients correspondants au même niveau d'échelle et à la même analyse angulaire.

Pour un pixel p de l'image f , si on constate que le coefficient qui lui correspond $c(n, o)$, pour une échelle n et un angle o , est significatif, alors on sauvegarde l'orientation o dans une liste l d'orientations associée au pixel p . En reproduisant cette analyse sur toutes les orientations pour tous les pixels, on obtient une caractérisation angulaire de chaque pixel du contour d'un objet.

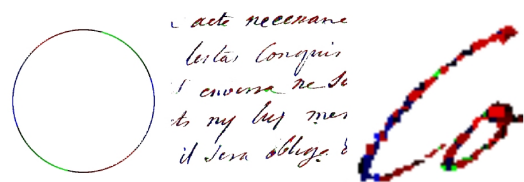


FIG. 1 – Évaluation des orientations sur un extrait de manuscrit autographe de Montesquieu, 1750

¹Le flux géométrique : champ indiquant les directions où les variations d'intensités sont régulières.

En s'inspirant des travaux de Jean-Pierre Antoine et Laurent Jacques dans [7], nous avons élaboré une évaluation de la courbure en fonction des orientations relevées précédemment. L'hypothèse formulée est la suivante : pour un pixel p d'une courbe L , plus le nombre d'orientations considérées comme significatives pour le pixel p est important, plus la courbure de L est élevée en ce point ; le cas extrême étant atteint dans le cas d'un pixel isolé.

Dans notre cas il nous suffit de récupérer la longueur de notre liste d'orientations l pour avoir une évaluation de la courbure en p .

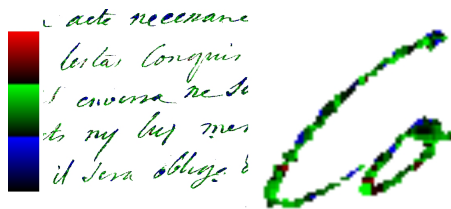


FIG. 2 – Évaluation des courbures sur un extrait de manuscrit autographe de Montesquieu

3.3 Calcul de la signature

Les deux primitives retenues pour le calcul de la signature des extraits de manuscrits sont la courbure et l'orientation sur les contours des formes d'une image. Le calcul de la signature est le suivant. Il s'agit d'une signature 2D avec en abscisse une échelle de la courbure et en ordonnée une échelle des orientations. Pour chaque pixel p on parcourt sa liste d'orientations l . Pour chaque élément o de l on forme un couple (c, o) , c étant la courbure correspondante à p et donc la longueur de l . On incrémente ensuite la valeur correspondante à ce couple dans la signature.



FIG. 3 – Signature de manuscrit médiéval

La mesure de distance se fait par le biais d'une corrélation linéaire entre les signatures considérées. Cette mesure de similarité $S(X, Y)$ entre deux signatures X et Y est définie comme suit :

$$S(X, Y) = \frac{Cov(X, Y)}{\sigma_x \sigma_y} \quad (3)$$

où $Cov(X, Y)$ est la covariance entre X et Y et σ_x et σ_y

les écarts-types.

Le principe consiste donc à interroger la base d'images (constituée dans notre cas de 2 corpus numérisés : un corpus médiéval et un corpus humaniste) à partir d'une requête image pour laquelle on veut trouver et trier toutes les images de la base qui lui sont similaires. Pour cela, les images sont décrites par une signature unique et caractéristique de l'écriture à partir d'une analyse des coefficients de curvelets.

4 Résultats et discussion

Dans l'optique d'une validation préliminaire de notre approche, nous avons élaboré un système de recherche d'images par le contenu. Ce système suit le schéma « classique » utilisé dans ce domaine ; il est illustré à la figure 4.

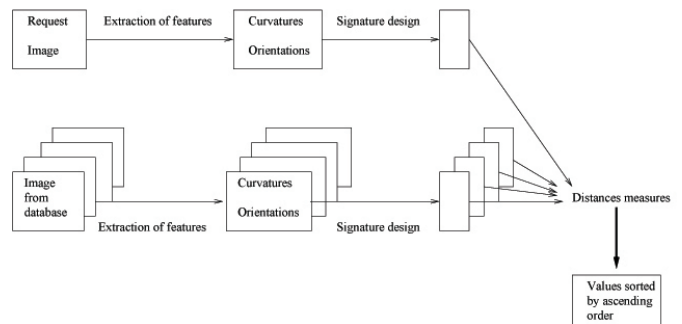


FIG. 4 – Schéma synaptique de recherche d'information dans le cas de bases d'images

Le corpus médiéval est constitué de manuscrits de la British Library (<http://www.bl.uk>) et de manuscrits de Digital Scriptorium (<http://sunsite.berkeley.edu>).

Sur le corpus humaniste, en moyenne, sur les 15 premières réponses de notre système, 74,85% correspondent à des images de l'auteur recherché. La courbe de Rappel/Précision est fournie en figure 5.

À partir des résultats de Rappel/Précision calculés pour cette courbe nous avons évalué une F-mesure ($F=2PR/(P+R)$) qui dans le cas de la base humaniste vaut 81,65 sur les 15 premières valeurs. Dans le cas de la base médiévale le résultat est de 43,41. Ce résultat bien inférieur au cas précédent est essentiellement dû aux origines différentes de nos images. En effet, l'étiquetage fait par les paléographes n'est pas le même selon les origines de ces derniers. Nous ne disposons donc pas de vérité terrain sur la classification des images médiévales et ne pouvons donc pas vérifier de manière systématique si une image renvoyée par notre système est réellement une erreur de classification ou simplement une différence d'étiquetage. Nous travaillons actuellement sur la construction d'une vérité terrain à partir de considérations purement perceptuelles d'une part et historiques d'autre part avec l'aide de paléographes de l'IRHT².

La deuxième évaluation que nous avons menée sur le même système concerne la robustesse de notre approche

²Institut de Recherche et d'Histoire des Textes

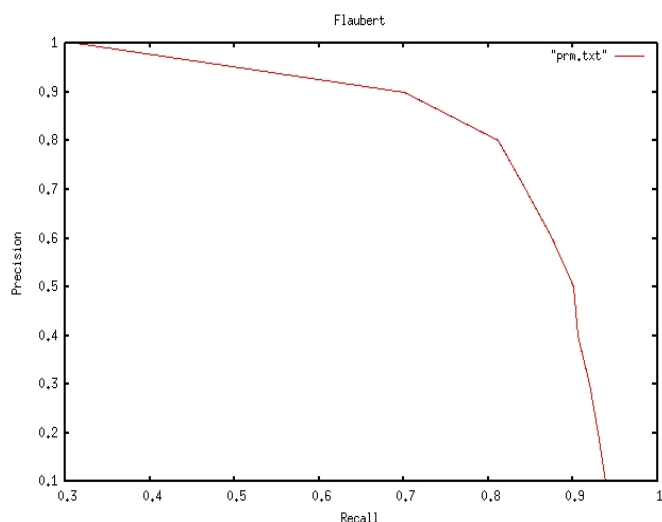


FIG. 5 – Courbe de Rappel/Précision sur la base humaniste

au bruit. Emmanuel Candès et Dave Donoho ont montré dans [4] que la transformée en Curvelets est bien adaptée au nettoyage d'images mais nous avons cherché à savoir si notre signature était elle aussi robuste en environnement bruité, ce qui est le cas de la plupart de nos images de manuscrits du patrimoine. L'analyse de robustesse porte sur les points suivants : changement de résolution, changement d'échelle, zooms sur des parties du texte, bruit Gaussien, duplication de mots sur la page, dilatations et érosions.

Sur les 33 images créées par ces dégradations, les 25 premières réponses de notre système ont un taux de réponse supérieur à 90%. Ces premières conclusions sont très satisfaisantes. Elles montrent qu'un outil d'analyse multi-résolution, principalement dédié à des applications de débruitage et de codage des images, et a priori non destiné à des applications de caractérisation et de représentation des contenus, a un réel potentiel discriminant dans les images de traits prises dans des conditions de numérisation non idéales (faible résolution à l'acquisition, image ayant subi des compressions avec pertes, ...)

5 Conclusion

Notre approche se veut l'exploitation d'une représentation multi-échelle alliant les informations de directions et de courbures pour l'extraction de primitives pertinentes aux frontières des formes. Notre choix s'est porté sur la transformée en Curvelets qui présente un potentiel très intéressant pour la caractérisation et l'analyse par le contenu. Les résultats portant sur un système de type CBIR appliqué à la recherche par le contenu dans les images de manuscrits numérisés sont très encourageants.

Références

- [1] N.G. Kingsbury, « Complex Wavelets for Shift Invariant Analysis and Filtering of Signals », *J. Applied Computer Harmonic Analysis*, vol. 10, no.3, pp.234-253, 2001.
- [2] M.N. Do et M. Vetterli, « The contourlet transform : an efficient directional multiresolution image representation », *IEEE Transactions on Image Processing*, vol. 14, no. 12, pp. 2091-2106, 2005.
- [3] E. Le Pennec et S.G. Mallat, « Sparse geometric image representation with bandelets », *IEEE Transactions on Image Processing*, vol. 14, no. 4, 2005.
- [4] E. Candès et D. Donoho, « Curvelets - A surprisingly effective nonadaptive representation for objects with edges, curves and surfaces », *Curves and Surfaces*, 1999. Vanderbilt University Press, Nashville, TN.
- [5] E. Candès et D. Donoho, « Recovering edges ill-posed inverse problems : optimality of Curvelet », *annals of statistics*, vol. 30, no. 3, pp. 784-842, 2002.
- [6] E. Candès et D. Donoho, « Ridgelets : A key to higher-dimensional intermittency ? », *Philosophical transactions - Royal Society. Mathematical, physical and engineering sciences*, vol. 357, no. 1760, pp.2495-2509, 1999.
- [7] J-P. Antoine et L. Jacques, « Measuring a curvature radius with directional wavelets », *Inst. Phys. Conf. Series*, vol. 173, pp. 899-904, 2003.
- [8] J. Radon, « Uber die Bestimmung von Funktionen durch ihre Integralwerte langs bestimmter Mannigfaltigkeiten », *Ber. Verb. Sachs. Akad. Wiss. Leipzig, Math.-Nat. Kl.*, vol. 69, pp. 262-277, 1917.