

Cartographie des échelles d'une image satellitaire

Bin Luo¹, Jean-François Aujol², Yann Gousseau¹, Henri Maître¹,

¹GET/Télécom Paris, CNRS UMR 5141, CNES-DLR-ENST Competence Centre

²CMLA, ENS Cachan, CNRS, PRES UniverSud

{bin-luo, gousseau, maitre}@tsi.enst.fr

Jean-Francois.Aujol@cmla.ens-cachan.fr

Résumé – Dans cet article, nous proposons une méthode qui permet d'associer une échelle caractéristique à chaque pixel d'une image. La définition de l'échelle caractéristique repose sur un filtrage de la carte topographique de l'image, constituée de l'ensemble de ses lignes de niveau. L'échelle d'un pixel est définie à partir de la forme (composante connexe d'une ligne de niveau) la plus contrastée le contenant. La carte des échelles ainsi obtenue a la particularité d'être spatialement très précise, ce qui est illustré sur des images satellitaires.

Abstract – In this article, a method for the estimation of a characteristic scale at each pixel of an image is proposed. The definition of characteristic scale relies on the topographic map of the image, made of the set of its level lines. The scale of a pixel is defined through the most contrasted shape (connected component of a level line) containing it. The resulting scale map is spatially very accurate. This fact is illustrated on several satellite images.

1 Introduction et motivations

De nombreux travaux traitent de l'extraction d'échelles caractéristiques à partir d'images numériques. Ces échelles sont liées aux tailles des structures (objets) constituant l'image. L'approche la plus classique, introduite par Lindeberg, consiste à étudier les extrema d'opérateurs différentiels dans un espace-échelle linéaire, [6]. De nombreuses méthodes alternatives ont été proposées, voir par exemple [1, 4, 16]. L'extraction d'une échelle caractéristique constitue une brique élémentaire de la plupart des méthodes de reconnaissance de structure, [8, 11], et est également un attribut efficace en discrimination ou classification d'image, [1, 7, 9]. En particulier les travaux présentés dans [7] démontrent la pertinence de l'échelle pour l'extraction de zones urbaines sur des images satellitaires.

Dans cette communication, nous proposons une méthode permettant de définir une échelle caractéristique en chaque pixel d'une image. Intuitivement, l'échelle d'un pixel est définie comme la taille de l'objet le plus contrasté contenant ce pixel. Formellement, l'approche proposée repose sur la carte topographique, constituée des frontières des ensembles de niveau de l'image. Il s'agit d'une représentation exhaustive, hiérarchique et géométrique du contenu de l'image, [2, 12]. Un filtrage de la carte topographique adapté aux images satellitaires est proposé. Nous exploitons la structure d'arbre de la carte topographique [13] pour nous affranchir des différentes étapes de lissage linéaire requises par les approches classiques de définition de l'échelle caractéristique, ce qui nous permet de définir une carte d'échelle spatialement précise. L'approche est enfin validée sur des images satellitaires.

2 Carte topographique d'une image numérique

La carte topographique [2] d'une image u est constituée des frontières topologiques de ses ensembles de niveau supérieurs :

$$\Psi_\lambda = \{x \in \mathbb{R}^2, u(x) \geq \lambda\}$$

pour $\lambda \in \mathbb{R}$ (les ensembles de niveau inférieurs pourraient de façon équivalente être considérés). Ces frontières sont appelées *lignes de niveau*. Une transformation rapide permettant de passer de l'image numérique à l'ensemble de ses lignes, la FLST (pour Fast Level Set Transform) a été proposée dans [14]. Il s'agit d'une représentation de l'image par un arbre dont les noeuds sont les composantes connexes des lignes de niveau. Ces composantes connexes sont des courbes de Jordan (courbes fermées et qui ne se croisent pas), dont l'intérieur est appelé une *forme*. Ces formes peuvent de façon équivalente être définies à partir des composantes connexes des ensembles de niveau. Pour chaque composante connexe A d'un ensemble de niveau (supérieur ou inférieur), on obtient une *forme* en considérant l'union de A et de ses trous (composantes connexes du complémentaire de A qui ne touchent pas le bord de l'image). Remarquons par ailleurs que l'image u peut être reconstruite à partir de l'arbre des formes en utilisant la relation :

$$u(x) = \sup\{\lambda | x \in \Psi_\lambda\}. \quad (1)$$

Outre son caractère géométrique, deux propriétés importantes de cette représentation sont les invariances par changements de contraste locaux et par l'opération d'inversion de contraste $u \rightarrow -u$. Cette dernière propriété implique que les objets clairs et sombres sont traités de manière équivalente, ce qui n'est pas le cas lorsque l'on tra-

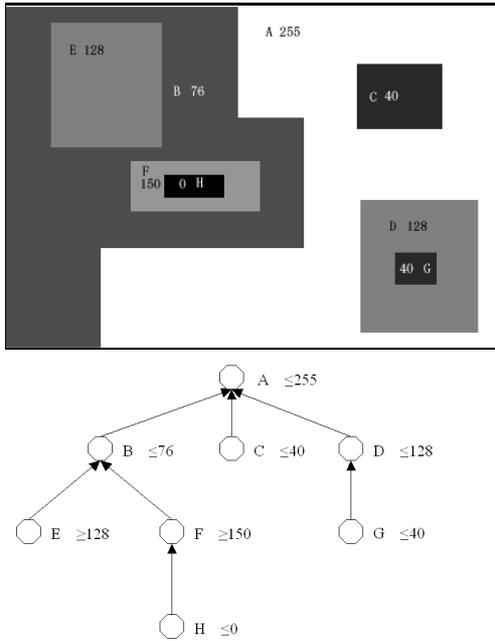


FIG. 1 – Exemple de FLST. Haut : image synthétique ; bas : arbre des formes obtenu par la FLST.

vaillent directement sur les ensembles de niveau, par exemple lors de l'utilisation des opérations les plus classiques de la morphologie mathématique.

Un arbre obtenu par FLST sur une image synthétique est montré Figure 1.

L'idée de départ de cette communication est d'associer à chaque pixel de l'image la forme (au sens noeud de la FLST) la plus pertinente parmi les formes contenant ce pixel, puis d'utiliser cette forme pour calculer une échelle caractéristique. Le paragraphe suivant détaille le filtrage de l'arbre proposé pour cette tâche.

3 Cartographie des échelles d'une image

Dans ce paragraphe, nous définissons un filtrage de l'arbre des formes d'une image permettant d'associer une forme à chaque pixel. Une notion de forme significative définie à partir de la carte topographique a été introduite dans [3]. La méthode repose sur une sélection des formes de l'image suffisamment contrastées. Cette sélection, dite *a contrario*, ne retient que les formes de l'image que leur périmètre et leur contraste rendent improbables dans une image de bruit et permet d'obtenir une représentation compacte et fidèle du contenu géométrique d'une image. Cette méthode, et ce n'est pas son objectif, n'associe donc pas ou peu de formes aux zones texturées de l'image. Pour le but que nous nous sommes fixé, associer à chaque point une échelle caractéristique, nous proposons un filtrage alternatif très simple de la carte topographique. D'une part ce filtrage associe bien une forme à chaque point et, d'autre part, il ne repose que sur un critère de contraste (et pas sur un critère géométrique) de manière à pouvoir associer n'importe quelle échelle à un point. En résumé, nous cher-

chons à associer à chaque pixel la forme la plus contrastée le contenant.

Contraste cumulé Pour un pixel x d'une image u , notons $\{f_i(x)\}_{i \in A(x)}$ l'ensemble des formes qui contiennent x , $A(x)$ étant un ensemble d'indices tels que $f_i(x) \subset f_{i+1}(x)$ (en remarquant que l'ensemble des formes contenant un pixel donné est totalement ordonné par inclusion). Pour chaque forme f_i , définissons $S(f_i)$ la surface de la forme, $P(f_i)$ son périmètre et $I(f_i)$ le niveau de gris associé à f_i . L'idée la plus simple pour associer un contraste à chaque forme consiste à définir le contraste d'une forme f_i comme la valeur absolue de la différence entre le niveau de gris de f_i et celui de f_{i-1} , soit

$$C(f_i) = |I(f_i) - I(f_{i-1})|.$$

Cependant, cette notion de contraste n'est satisfaisante que si l'on suppose qu'à chaque objet est associée une unique forme. Or, dans une image naturelle numérique, les contours des objets sont toujours lissés en raison des filtrages dus à la formation de l'image. L'image étant quantifiée, ceci se manifeste sous la forme de paquets de formes emboîtées, conduisant le plus souvent à une différence de niveau de gris de 1 entre les formes. Une solution à ce problème peut être de calculer le contraste d'une ligne dans un voisinage et de sélectionner la ligne la plus pertinente le long de branches monotones de l'arbre (voir [3]). Nous intéressés aux images satellitaires pour lesquelles le flou est uniforme sur l'image et connu (nous reviendrons plus loin sur ce point), nous choisissons pour pallier ce problème de cumuler le contraste (valeur absolue de la différence de niveau de gris entre deux formes consécutives) des formes suffisamment proches en utilisant la structure de la carte topographique. Pour ce faire, sont cumulés les contrastes des formes f_i telles que

$$S(f_i) - S(f_{i-1}) \leq \lambda P(f_{i-1}), \quad (2)$$

où λ est une constante (un critère semblable ne reposant que sur la surface des formes a été proposé dans [12]). Ce critère repose sur l'hypothèse que les lignes de niveau correspondant à un contour flou sont régulièrement espacées d'une distance λ . Nous définissons donc le contraste cumulé d'une forme f_i par la formule

$$\bar{C}(f_i) = \sum_{k=a(i)}^i C(f_k),$$

où pour tout i ,

$$a(i) = \min\{j | \forall k = j+1, \dots, i, S(f_k) - S(f_{k-1}) \leq \lambda P(f_{k-1})\}$$

Si le nombre $a(i)$ n'est pas défini (c'est-à-dire si la relation (2) n'est pas vérifiée), alors $\bar{C}(f_i) = C(f_i)$. Le contraste cumulé d'une forme f_i est donc obtenu en ajoutant les contrastes de lignes de niveau suffisamment proches, correspondant généralement au même contour.

Lien entre contraste cumulé et flou Le contraste cumulé est défini en supposant que les lignes de niveau correspondant à un même contour sont espacées d'une distance λ . Plus précisément, on suppose que $f_{i+1} = f_i \oplus D(\lambda)$, où \oplus est l'addition de Minkowski¹ et $D(\lambda)$ le disque

¹Pour deux ensembles A et B , $A \oplus B = \{x + y, x \in A, y \in B\}$.

centré à l'origine de rayon λ . Ceci constitue une approximation de l'effet du flou sur les lignes. Dans ce cas, la formule de Steiner [15] s'écrit

$$S(f_{i+1}) = S(f_i) + \pi\lambda^2 + \lambda P(f_i), \quad (3)$$

ce qui justifie, au premier ordre, l'utilisation de la formule (2). Par ailleurs, supposer que les lignes sont équidistantes revient en première approximation à supposer que le noyau de flou est une fonction indicatrice. Si le noyau a une forme plus complexe (e.g. gaussien) l'écart entre lignes variera selon leur position par rapport au contour. Remarquons également que cette distance interlignes dépend du contraste du contour, de la quantité de flou et de la quantification de l'image. On peut montrer facilement qu'en supposant un contraste minimum des contours et en connaissant la largeur du noyau de flou, il est possible de déduire une borne supérieure sur la valeur de λ à utiliser. Remarquons également que l'hypothèse d'un noyau de flou constant sur l'image et connu est réaliste dans le cas des images satellitaires mais l'est beaucoup moins dans le cas d'images naturelles pour lesquelles le flou associé à un objet dépend de sa position, voir [5].

Forme la plus contrastée associée à un pixel La forme la plus contrastée contenant le pixel x est définie par

$$\hat{f}(x) = f_{\arg \max_{i \in A(x)} \bar{C}(f_i)}. \quad (4)$$

Dans le cas où le max est atteint en plusieurs indices, l'indice le plus petit est retenu. Enfin, nous choisissons de considérer comme forme finale associée à x la forme $\hat{f}(x)$ à laquelle on retire les formes les plus contrastées qu'elle contient. Rappelons en effet qu'une forme est une composante connexe d'un ensemble de niveau dont les trous (composantes connexes du complémentaire ne touchant pas le bord de l'image) ont été bouchés. Ainsi, sur la Figure 1, la forme F contient les pixels de la forme H . Avec en vue une représentation fidèle du réseau routier, ingrédient important dans les images satellitaire, nous définissons la forme la plus contrastée contenant x par

$$\tilde{f}(x) = \hat{f}(x) \setminus \bigcup_{\hat{f}(y) \subsetneq \hat{f}(x)} \hat{f}(y),$$

c'est-à-dire la forme $\hat{f}(x)$ à laquelle ont été retirées les formes les plus contrastées strictement incluses dans $\hat{f}(x)$. D'autres choix sont bien entendu possible selon l'application visée (en particulier celui de conserver la forme d'origine $\hat{f}(x)$). Sur la figure 2 sont affichées les formes les plus contrastées d'une image satellitaire de résolution 5m (remarquons que les formes $\tilde{f}(x)$ forment une partition de l'image). Pour l'affichage, le niveau de gris de chaque forme est le niveau de gris moyen des pixels de cette forme.

Echelle caractéristique locale Afin d'associer une échelle caractéristique à chaque pixel x , nous considérons la forme $\tilde{f}(x)$. La première notion d'échelle à laquelle on peut penser est $\sqrt{S(\tilde{f}(x))}$. Cette définition donne des résultats satisfaisant, sauf en présence de structures fines et étendues,

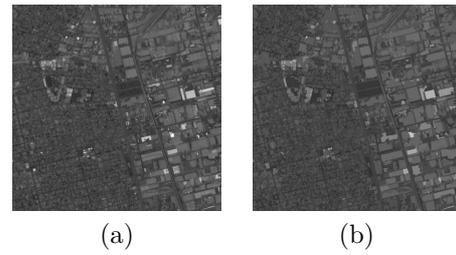


FIG. 2 – (a) Image de Los Angeles (SPOT5 5m); (b) Formes les plus contrastées correspondantes.

tels à nouveau les réseaux routiers. Nous lui préférons une échelle définie comme le rapport surface/périmètre. L'échelle caractéristique d'un pixel x est par conséquent définie par

$$e(x) = \frac{S(\tilde{f}(x))}{P(\tilde{f}(x))}.$$

D'autres choix sont bien entendu possibles, tel que le rayon du plus grand disque inclus dans la forme.

4 Expériences

Les expériences numériques sont effectuées en utilisant l'implantation de la FLST disponible dans l'environnement de programmation libre *Megawave*². Nous montrons les cartes d'échelles obtenues sur deux images aériennes SPOT5 (©CNES), de résolution 5m, prises respectivement à Los Angeles et Toulouse et sur une image de QuickBird, de résolution 0.61m, prise à Ouagadougou. Pour choisir la valeur de λ à utiliser, nous n'avons pas utilisé de connaissance a priori du noyau de flou. Partant de l'observation empirique que, sur les images satellitaires considérées, les lignes correspondant au même contour sont espacées d'environ 1 pixel, nous utilisons la valeur $\lambda = 1.2$, de manière à prendre en compte le terme quadratique de la formule de Steiner (3) et la variabilité de l'image. Par ailleurs, pour l'image de Ouagadougou, seules les formes de surface supérieure à 16 pixels sont prises en compte, la haute résolution de cette image donnant une carte d'échelle plus bruitée que sur les images de résolution 5m. Ce sont là les deux seuls paramètres de la méthode proposée.

On observe que les échelles calculées sont en bonne adéquation avec la taille des structures présentes dans les différentes zones des images. Observons également que, la résolution des images étant connue et l'échelle calculée correspondant à la taille d'un élément géométrique, cette échelle est directement liée à la taille réelle des structures de l'image. Ceci ne peut se faire aussi précisément en utilisant les notions d'échelles liées à l'espace-échelle linéaire, même si des approximations en ce sens sont possibles, voir [10]. Enfin, et c'est là un point fort de l'approche proposée, remarquons que la carte des échelles est très précise spatialement. Ceci est du au fait que, contrairement aux approches les plus classiques d'évaluation de l'échelle caractéristique, l'échelle n'est pas calculée sur une fenêtre glissante. Le voisinage d'un pixel sur lequel l'échelle est

²<http://www.cmla.ens-cachan.fr/Cmla/Megawave/>

calculée dépend de la structure de la carte topographique et s'adapte donc à la structure de l'image.

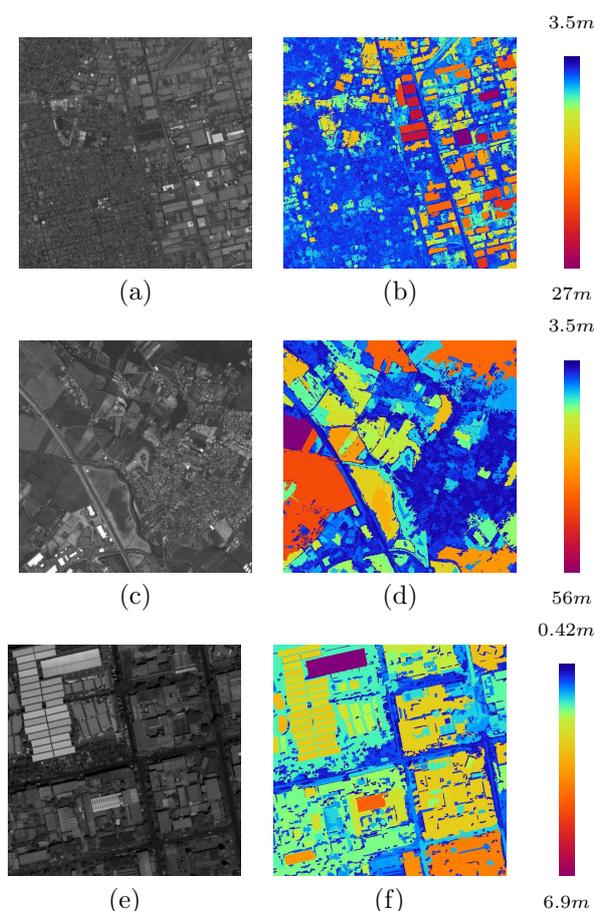


FIG. 3 – (a) Image de Los Angeles, SPOT5 (5m)©CNES; (b) Image d'échelle calculée à partir de l'image (a); (c) Image de Toulouse, SPOT5 (5m)©CNES; (d) Image d'échelle calculée à partir de l'image (c); (e) Image d'Ouagadougou, QuickBird (0.61m); (f) Image d'échelle calculée à partir de l'image (e), les formes dont les surfaces sont inférieures que 16 pixels ne sont pas prise en compte.

5 Conclusions et perspectives

Nous avons présenté une méthode permettant de définir une échelle caractéristique en chaque pixel d'une image en utilisant la carte topographique. L'échelle d'un pixel est définie comme l'échelle de la forme la plus contrastée contenant ce pixel. Les expériences montrent que notre méthode donne des résultats spatialement très précis.

Pour exploiter la carte d'échelle obtenue en indexation ou classification d'images, nous comptons utiliser un processus de régularisation. Dans cette perspective, la carte d'échelle devient une primitive géométrique de l'image (nous pensons qu'elle se révélera particulièrement utile pour indexer les zones urbaines dans les images satellitaires). Des expériences numériques sont en cours pour valider cette approche, ainsi que des comparaisons avec les méthodes existantes, notamment [7].

Références

- [1] Brox, T. et J. Weickert: *A TV flow based local scale measure for texture discrimination*. Dans *ECCV 04*, tome 2, pages 578–590, Mai 2004.
- [2] Caselles, V., B. Coll et J.M. Morel: *Scale Space Versus Topographic Map for Natural Images*. Dans *Scale-Space*, pages 29–49, 1997.
- [3] Desolneux, A., L. Moisan et J.M. Morel: *Edge Detection by Helmholtz Principle*. *Int. J. of Computer Vision*, 14(3) :271–284, 2001.
- [4] Kadir, T. et M. Brady: *Scale, Saliency and Image description*. *Int. J. of Computer Vision*, 45(2) :83–105, 2001.
- [5] Ladjal, S.: *Flou et quantification dans les images numériques*. Thèse de doctorat, ENS de Cachan, France, 2005.
- [6] Lindeberg, T.: *Feature Detection with Automatic Scale Selection*. *Int. J. of Computer Vision*, 30 :79–116, 1998.
- [7] Lorette, A.: *Analyse de Texture par Méthodes Markoviennes et par Morphologie Mathématique : Application à l'Analyse des Zones Urbaines sur des Images Satellitaires*. Thèse de doctorat, Université de Nice, 1999.
- [8] Lowe, D.G.: *Distinctive image features from scale-invariant keypoints*. *Int. J. of Computer Vision*, 60(2) :91–110, 2004.
- [9] Luo, B., J F. Aujol, Y. Gousseau, S. Ladjal et H. Maître: *Characteristic scale in satellite images*. Dans *ICASSP 2006*, tome 2, pages 809–812, 2006.
- [10] Luo, B., J F. Aujol, Y. Gousseau, S. Ladjal et H. Maître: *Resolution independent characteristic scale with application to satellite images*. Internal publication of ENST Paris, 2006.
- [11] Mikolajczyk, K. et C. Schmid: *Scale and affine invariant interest point detectors*. *Int. J. of Computer Vision*, 60(1) :63–86, 2004.
- [12] Monasse, P.: *Morphological representation of digital images and application to registration*. Thèse de doctorat, Paris IX, 2000.
- [13] Monasse, P. et F. Guichard: *Scale-Space from a Level Lines Tree*. Dans *Scale-Space'99*, pages 175–186, 1999.
- [14] Monasse, P. et F. Guichard: *Fast Computation of a Contrast Invariant Image Representation*. *IEEE Transactions on Image Processing*, 9(5) :860–872, Mai 2000.
- [15] Stoyan, D., W. S. Kendall et J. Mecke: *Stochastic Geometry and its Applications*. Wiley, 2nd édition, 1995.
- [16] Winter, A., H. Maître, N. Cambou et E. Legrand: *An Original Multi-Sensor Approach to Scale-Based Image Analysis for Aerial and Satellite Images*. Dans *IEEE-ICIP-97*, tome II, pages 234–237, Santa Barbara, CA, USA, 1997.