HAL
open science

Segmentation pyramidale localisée dans un ruban fermé

Pascal Bertolino, Jérémy Huart, Guillaume Foret

To cite this version:

Pascal Bertolino, Jérémy Huart, Guillaume Foret. Segmentation pyramidale localisée dans un ruban fermé. 20ème Colloque GRETSI, 2005, Louvain-la-Neuve, Belgique. pp.CD. hal-00372194

HAL Id: hal-00372194

https://hal.science/hal-00372194v1

Submitted on 7 May 2009

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.

Segmentation pyramidale localisée dans un ruban fermé

Pascal BERTOLINO[1], Jérémy HUART[1], Guillaume FORET[2]

[1]Laboratoire des Images et des Signaux,
BP 46, 38402 Saint Martin d'Hères, France

[2]Equipes Traitement des Images et du Signal,
BP 44, 95014 Cergy-Pontoise, France
pascal.bertolino@inpg.fr, jeremy.huart@inpg.fr

Résumé – Nous montrons comment un processus de segmentation peut être utilisé localement sur des régions d'intérêt d'une image, délimitées grossièrement par l'utilisateur. L'algorithme de la pyramide irrégulière convient particulièrement bien à cette approche : une propagation concurrente d'étiquettes de *racines* dans un graphe de régions permet de classifier les pixels à l'intérieur ou à l'extérieur de l'objet considéré.

Abstract – This paper shows how a segmentation process can be locally used within regions of interest roughly delineated by the user, in a still image. The graph pyramid (aka irregular pyramid) algorithm is used to perform a concurrent root label propagation that classifies each pixel inside or outside a given object.

1 Introduction

La segmentation d'images naturelles précise et efficace peut rarement être effectuée automatiquement ; c'est pourquoi le recourt à la segmentation interactive [1, 2] est souvent nécessaire dans de nombreuses applications : son but est d'extraire un ou plusieurs objets d'intérêt en proposant à l'utilisateur potentiel un outil simple et pratique à utiliser fournissant des résultats facilement exploitables.

Les contours actifs [3, 4] sont des méthodes très classiques qui nécessitent une initialisation avec un contour grossier à l'intérieur ou à l'extérieur de l'objet. Malheureusement, les snakes sont généralement très sensibles aux minima locaux et assez difficiles à paramétrer. La ligne de partage des eaux [5] quant-à-elle, fournit une sur-segmentation de l'image. Dans [6], les auteurs proposent un traitement interactif où les marqueurs sont sélectionnés par des clics avec la souris. Dans [1], la LPE fournit une segmentation multi-échelle dans laquelle l'utilisateur peut sélectionner des régions d'intérêt. Dans [7], après un seuillage automatique, les contours des régions sont représentés par des courbes paramétriques. Leurs points de contour peuvent être édités par l'utilisateur. Les ciseaux intelligents ont été développés par [2] pour la composition. La méthode détecte un contour d'objet potentiel proche du pointeur de souris et trouve le meilleur chemin partant d'un pixel de départ.

Dans notre approche, une croissance de régions est appliquée uniquement près du bord de l'objet indiqué grossièrement par l'utilisateur.

2 Pyramide d'image irrégulière *vs* pyramide locale

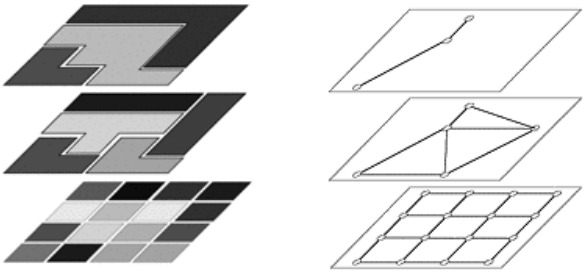
La pyramide irrégulière [8] [9] est un outil de segmentation puissant. Le principe de cette méthode est d'initialiser un graphe d'adjacence, où chaque sommet correspond à une région monopixelique de l'image. Utilisant un algorithme de type parallèle, les régions voisines similaires peuvent fusionner, entraînant une décroissance du nombre de sommets (figure 1.a). Les régions i et j sont similaires ($R_i \sim R_j$) si la distance entre leur couleur moyenne, dans l'espace YUV, est plus faible qu'un certain seuil : $R_i \sim R_j$ si $d(YUV(R_i),YUV(R_j)) < T$. Le traitement, qui porte sur toute l'image, est itéré jusqu'à stabilisation du nombre de régions.

Dans la pyramide locale que nous proposons, seul un nombre réduit de pixels de l'image sont identifiés à des sommets, tandis que les n composantes connexes restantes sont identifiées à n sommets (figure 1.b et c), et sont des *racines* (régions qui appartiendront au résultat final) tel que le fond par exemple. Cette focalisation sur une zone particulière de l'image nécessite une interaction humaine, et peut être réalisée à la souris en marquant grossièrement le contour de la zone à l'aide d'un ruban fermé (figure 2.a).

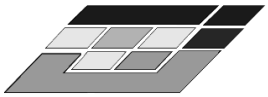
3 Propagation des étiquettes

Une fois dessiné, le ruban induit trois composantes connexes, soit trois étiquettes : l'extérieur, l'intérieur et le ruban lui-même, supposé contenir le vrai contour de l'objet (figure 2.b). Les pixels de l'extérieur appartiennent au fond (le premier sommet racine) tandis que ceux de l'intérieur appartiennent à l'objet d'intérêt (un second sommet racine). Tous les pixels formant le ruban représentent une zone indéfinie. Ces derniers seront segmentés afin de fusionner avec l'une ou l'autre des racines.

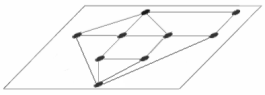
A la base de la pyramide locale, le graphe d'adjacence inclut un sommet pour chaque racine et un sommet pour chaque pixel de la zone indéfinie (figure 1.c). A chaque racine est associée une étiquette unique. Les sommets (i.e. pixels) de la zone indéfinie restent non étiquetés. Une fine couche de pixels de part et d'autre du ruban sont séparés des racines et constituent autant de sommets portant l'étiquette de la racine correspondante (figure 2.d). Ce sont ces deux couches qui vont permettre



(a) Exemple de pyramide de graphe construite sur une image 4 $\times$ 4 : empilements de partitions et de graphes

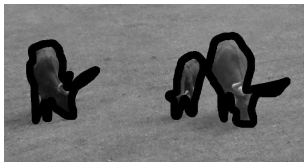


(b) Exemple d'une initialisation de pyramide locale



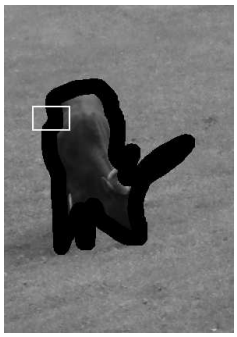
(c) Graphe correspondant à (b)

FIG. 1 – *Pyramide d'image vs pyramide locale*

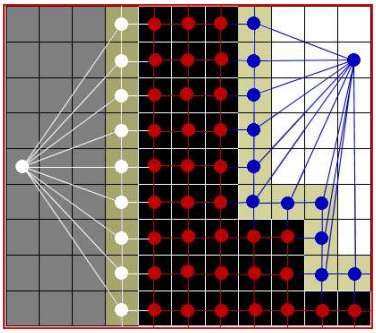


(a) Initialisation de régions d'intérêt

(b) Etiquetage initial des racines induites par (a). Les zones noires n'ont pas d'étiquette



(c) Le rectangle est zoomé dans (d)



(d) Graphe d'adjacence de la zone marquée dans (c)

FIG. 2 – *De l'initialisation manuelle d'une zone d'intérêt à la définition d'un graphe d'adjacence où sera effectuée la segmentation*

(en fusionnant avec d'autres pixels) de propager les étiquettes des deux racines dans la zone indéfinie. Durant la construction de la pyramide locale, les fusions s'effectuent toujours selon le critère de similarité, mais respectent également des règles qui assurent la propagation cohérente des étiquettes à travers la zone indéfinie.

Le regroupement de sommets [8] se fait localement selon une phase de décimation (des sommets disparaissent) et une phase de rattachement (les pixels des sommets disparus se rattachent au sommet voisin survivant le plus similaire). Le nombre de sommets se regroupant en un nouveau sommet n'est pas fixé *a priori*. Les phases de décimation et de rattachement nécessitent que la similarité soit évaluée entre tous couples de sommets. Si n sommets ($n \geq 2$) sont similaires, ils pourront éventuellement être fusionnés en une seule et même région, selon les phases décimation/regroupement. Ainsi, la notion de similarité est capitale ; rajouter de nouvelles contraintes dans les règles de construction de la pyramide peut simplement être vu comme modifier le critère de similarité. Pour prendre en compte ces nouvelles contraintes, on modifie le critère vu en section 2 et on rajoute des règles concernant la propagation des étiquettes. Par la suite, $R(l)$ est une région portant l'étiquette l et R est une région non étiquetée ; $\sim$ est le symbole de similarité.

- **Règle de similarité :** $R_i(l_i) \nsim R_j(l_j)$ si $l_i \neq l_j$ ou si $d(YUV(R_i),YUV(R_j)) \geq T$. Deux régions portant deux étiquettes différentes ne peuvent pas fusionner, ceci pour éviter la fusion de différents objets.
- **Règle de propagation 1 :** Si $R_i \sim R_j(l)$,$R_i \cup R_j(l) = R_k(l)$. Une région non étiquetée fusionnant avec une région étiquetée l donne naissance à une région portant l'étiquette l.
- **Règle de propagation 2 :** Si $R_i(l) \sim R_j(l)$,$R_i(l) \cup R_j(l) = R_k(l)$. La fusion de plusieurs régions d'étiquette l donne naissance à une région portant l'étiquette l.
- **Règle de propagation 3 :** Si $R_i \sim R_j$,$R_i \cup R_j = R_k$. Des régions non étiquetées fusionnant entre elles donnent naissance à une région non étiquetée.

Lorsque plus aucune fusion n'est possible et qu'il reste éventuellement des sommets non étiquetés, la règle de similarité est relâchée : $R_i(l_i) \nsim R_j(l_j)$ ssi $l_i \neq l_j$, $R_i(l_i) \sim R_j(l_j)$ sinon. La partition finale comporte alors autant d'objets que de racines.

4 Résultats

Il est naturel que la méthode fonctionne bien dans le cas d'objets homogènes sur un fond homogène.

Beaucoup plus intéressant, la méthode fonctionne bien dans l'hypothèse où les objets ont une colorimétrie hétérogène sur un fond relativement homogène (ou inversement), même lorsque le signe du gradient s'inverse à la frontière entre le fond et l'objet. Les dégradés sont aussi bien pris en compte. Cette approche est intéressante notamment lorsque les objets ont une forme complexe.

Les résultats présentés ne sont pas post-traités, et sont obtenus avec le même seuil de similarité T. La figure 3.b montre comment il est souvent impossible avec des méthodes simples (ici un seuillage) d'extraire correctement un objet, soit à cause de sa texture, de son ombre portée ou du manque de contraste localement. On peut remarquer que le contour du dos de l'animal (figure 3.a) n'est pas du tout contrasté, ni en couleurs, ni en niveaux de gris.

Les résultats (figures 3 et 4) montrent que notre méthode est peu sensible à la façon dont l'utilisateur initialise le traitement :

en effet, une variabilité importante de l'épaisseur du ruban, de son positionnement (plutôt intérieur ou extérieur), de sa forme (plus ou moins régulière), fournit des résultats très similaires. Le ruban peut être tracé soit à l'aide de segments (figures 3.c,e, 4.a,c), soit à main levée (figure 4.e,g).

Les figures 3.d et f présentent deux résultats obtenus avec deux épaisseurs différentes (10 et 40 pixels). Le ruban de la figure 3.e est obtenu avec 10 clics de la souris seulement ; en revanche, il fournit des détails précis de la tête et des jambes de l'animal.

Dans la figure 4.a, la segmentation est obtenue avec un ruban positionné plutôt dans l'objet alors que dans la figure 4.c, il est plutôt dans le fond. Bien que ces résultats ne soient pas exactement identiques, ils diffèrent assez peu.

Dans les figures 4.e et g, on a varié la forme du ruban : l'utilisateur peut effectivement dessiner des "pâtés" qui contiennent de nombreux détails de contour afin de les extraire correctement, comme le montre la jambe arrière dans la figure 4.g.

Enfin, la méthode présentée peut être utilisée telle quelle, sans modification, avec un nombre quelconque de racines et un ruban fermé peut avoir une topologie quelconque et être connexe à un nombre quelconque de racines. comme le montre l'exemple de la figure 5. L'approche proposée garantit un nombre d'objets final égal au nombre de racines. Il est donc important que le ruban soit fermé pour que l'intérieur puisse se distinguer de l'extérieur.

5 Discussion

Nous proposons une méthode bien adaptée aux images complexes où le contour d'un objet peut successivement prendre des configurations très différentes : gradient positif, négatif, contours multiples. La figure 6 montre un résultat obtenu avec notre application ExtraK'Obs[1] pour un objet qui se distingue assez difficilement du fond.

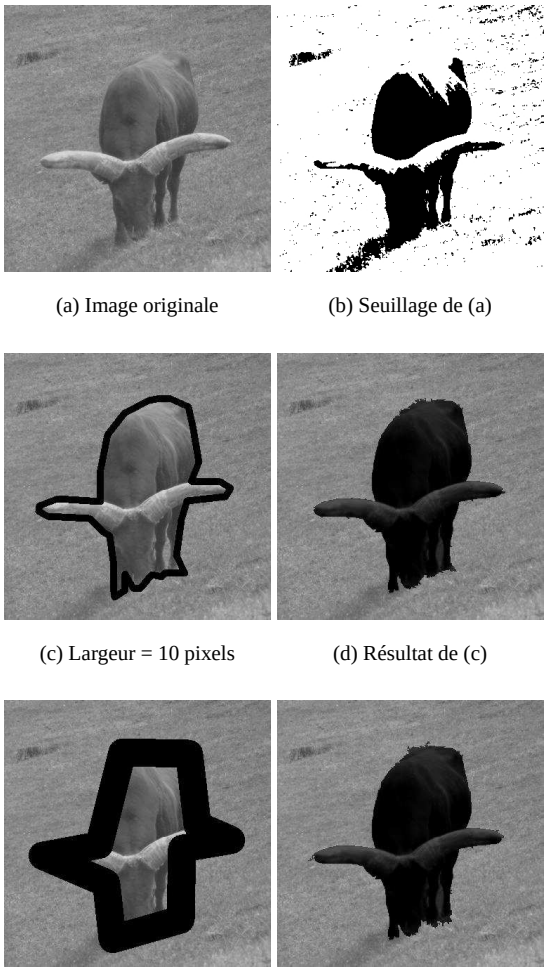
La texture du fond et/ou de l'objet est bien entendu un problème important auxquelles doivent faire face la majorité des méthodes de segmentation et notamment celle que nous présentons. La (re-)connaissance de présence de textures (objet ou fond) et de contours pourrait servir à favoriser certaines fusions lors de la propagation concurrente d'étiquettes. Il semblerait par exemple judicieux de favoriser la propagation dans les zones peu texturées qui pourraient alors s'étendre jusqu'à être stoppées par des zones plus texturées.

La localisation de la frontière finale n'est pas liée géométriquement à la localisation et à l'épaisseur du ruban. Ces deux paramètres peuvent néanmoins influencer le résultat final dans le sens où une localisation et une largeur mieux adaptées limiteront les erreurs de segmentation (faux contour, fuites, ...). La frontière finale correspond localement à la frontière entre deux régions qui ont crû puis acquis une étiquette. Ainsi la croissance de régions non encore étiquetées est un élément essentiel que nous devons étudier de manière plus approfondie.

6 Conclusion

Nous avons montré dans ce travail préliminaire une méthode associée à une interaction qui fournit à un utilisateur un outil permettant une extraction facile et précise d'objets. Maintenant, pour améliorer ces premiers résultats, nous envisageons deux pistes différentes : tout d'abord générer de façon automatique les zones d'incertitude par un pré-traitement des images. Ensuite, nous voulons intégrer d'autres critères permettant de régulariser localement les contours détectés pendant la segmentation.

1. $http://www.lis.inpg.fr/pages_perso/bertolino/software.php$



(a) Image originale (b) Seuillage de (a)

(c) Largeur = 10 pixels (d) Résultat de (c)

(e) Largeur = 40 pixels (f) Résultat de (e)

FIG. 3 – *Une variabilité importante de l'épaisseur du ruban entraîne une faible variabilité des résultats*

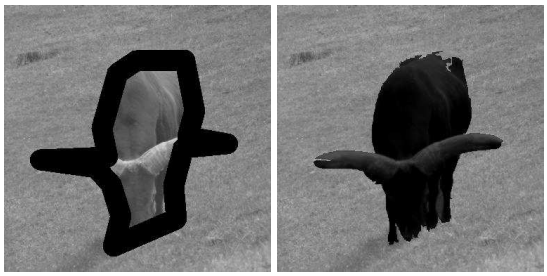
Références

[1] F. Zanoguera, B. Marcotegui, and F. Meyer, "A toolbox for interactive segmentation based on nested partitions," in *IEEE International Conference on Image Processing, ICIP'99*, Kobe, Japan, 1999.

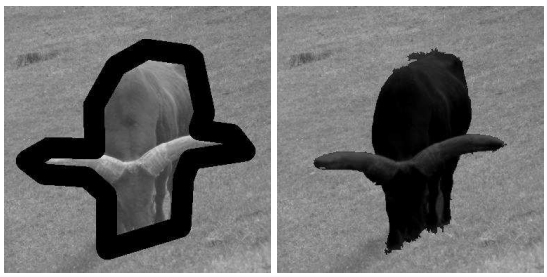
[2] Eric Mortensen and Wiliam A. Barrett, "Intelligent scissors for image composition," in *Proc. of the ACM SIGGRAPH 95 : Computer Graphics and Interactive Techniques*, Los Angeles, August 1995, pp. 191–198.

[3] M. Kass, A. Witkins, and D. Tersopoulos, "Snakes: Active contour models," *Computer Vision Graphics and Image Processing*, pp. 321–331, 1988.

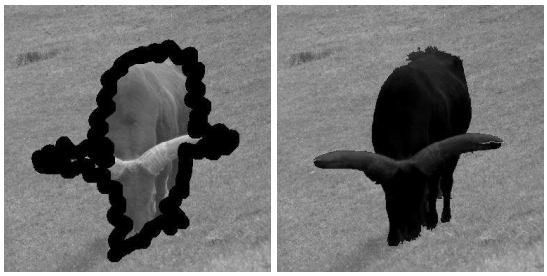
[4] A.K. Jain, Y. Zhong, and M.P. Dubuisson-Jolly, "Deformable template models: A review," *Signal Processing*, vol.



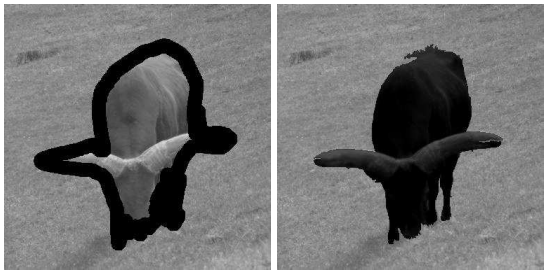
(a) Ruban intérieur (b) Résultat de (a)



(c) Ruban extérieur (d) Résultat de (c)



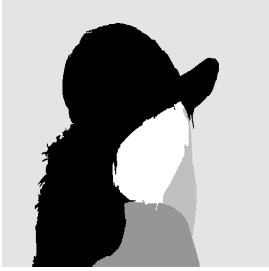
(e) Ruban bruité (f) Résultat de (e)



(g) Ruban irrégulier (h) Résultat de (g)

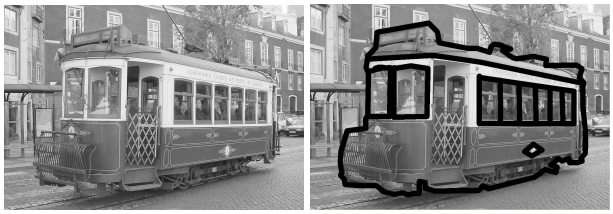
FIG. 4 – *Une variabilité importante du positionnement et de la régularité du ruban entraîne une faible variabilité des résultats*

(a) Image originale (b) Initialisation



(c) Partition obtenue

FIG. 5 – *Segmentation de plusieurs régions d'intérêt*



(a) Image originale (b) Initialisation

(c) Partition obtenue

FIG. 6 – *Exemple de résultat obtenu avec l'application Extra-K'Obs*

71, no. 2, pp. 109–129, 1998.

[5] S. Beucher and F. Meyer, "The morphological approach to segmentation: the watersheds transformation in e.r. doughterty (ed.)," in *Mathematical Morphology in Image Processing*, New-York: Dekker, 1992, pp. 433–481.

[6] J. Cutrona and N. Bonnet, "Two methods for semi-automatic segmentation based on fuzzy connectedness and watersheds," in *IASTED International Conference on Visualisation, Imaging and Image Processing VIIP' 2001*.

[7] M. Jackowski, A. Goshtasby, and M. Satter, "Interactive tools for image segmentation," in *SPIE's International Symposium on Medical Imaging*, San Diego, USA, 1999.

[8] A. Montanvert, P. Meer, and A. Rosenfeld, "Hierarchical image analysis using irregular tessellations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 4, pp. 307–316, April 1991.

[9] J.M. Jolion and A. Montanvert, "The adaptive pyramid a framework for 2d image analysis," *CVGIP Image Understanding*, vol. 55, no. 3, pp. 339–348, 1992.