



Open Archive TOULOUSE Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : <http://oatao.univ-toulouse.fr/>
Eprints ID : 18766

The contribution was presented at RFIA 2016 :
<http://rfia2016.iut-auvergne.com/>

To cite this version : Mathieu, Bérengère and Cruzil, Alain and Puel, Jean-Baptiste *Segmentation interactive pour l'annotation de photographies de paysages*. (2016) In: Congrès Francophone de Reconnaissance des Formes et Intelligence Artificielle (RFIA 2016), 27 June 2016 - 1 July 2016 (Clermont-Ferrand, France).

Any correspondence concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

Segmentation interactive pour l'annotation de photographies de paysages

Bérengère Mathieu^{1*}

Alain Crouzil¹

Jean-Baptiste Puel²

Université de Toulouse ; UPS¹ ; ENFA² ; IRIT, 118 route de Narbonne, F-31062 Toulouse, France

berengere.mathieu@irit.fr

Domaine principal de recherche : RFP

Papier soumis dans le cadre de la journée commune : NON

Résumé

Cet article traite de la segmentation interactive dans le cas de problèmes multiclassés. Nous proposons une nouvelle méthode capable de retrouver les objets sémantiques présents dans une image, à partir d'échantillons donnés par un utilisateur. Notre méthode a été testée avec succès sur deux problèmes de référence. Nous montrons par ailleurs qu'elle s'avère extrêmement utile pour réaliser des annotations de photographies.

Mots-clés

Vision par ordinateur, segmentation interactive, classification par SVM, superpixels.

Abstract

This paper addresses the problem of interactive multiclass segmentation. We propose a new method, which extracts relevant semantic objects from a few strokes drawn by a human user over an image. We demonstrate that SCIS significantly outperforms competing algorithms by evaluating its performances on two reference benchmarks. Moreover we show that this method is suitable for photograph annotation.

Keywords

Computer vision, interactive segmentation, SVM classification, superpixels.

1 Introduction

Aujourd'hui encore et malgré les nombreux travaux sur le sujet, segmenter une image, ou en d'autres termes, grouper ses pixels en régions correspondant aux éléments visibles, demeure un problème ouvert. L'une des difficultés les plus délicates à surmonter concerne l'incertitude quant au résultat attendu : telle segmentation jugée pertinente dans un premier contexte, peut se révéler inutilisable pour un autre type d'application. La figure 1 illustre ce problème : pour une même image, elle montre deux segmentations possibles.

La segmentation interactive contourne cette difficulté en demandant à l'utilisateur de fournir quelques informa-

tions (points de contours, échantillons de pixels pour chaque région, etc.). Ces indications permettent de déterminer les caractéristiques des différentes régions ainsi que leur nombre. Bien que la plupart des méthodes proposées se bornent à résoudre des problèmes à deux classes [1, 3, 4, 7, 11, 18], quelques méthodes récentes permettent à l'utilisateur d'extraire plusieurs classes [2, 16]. Cet article propose un nouvel algorithme de segmentation interactive, SCIS (« *Superpixel Classification-based Interactive Segmentation* »). Cette méthode simple permet d'extraire précisément un nombre arbitraire d'éléments, tout en conservant des temps de calcul largement inférieurs à ceux des méthodes existantes. Dans les sections suivantes, nous dresserons un état de l'art du domaine de la segmentation interactive. Puis, nous décrirons la méthode de segmentation interactive multiclassée que nous proposons. Nous évaluerons les performances de SCIS grâce à deux problèmes de référence : celui de McGuinness *et al.* [11] et celui de Santner *et al.* [16]. Enfin, nous terminerons en montrant l'utilité de notre méthode dans un domaine d'application particulier : l'annotation de photographies issues d'observatoires du paysage.

2 État de l'art

Les méthodes de segmentation interactives font partie des méthodes de segmentation dites semi-automatiques. L'utilisateur choisit quelques pixels (les germes) et indique pour chacun d'eux l'élément sémantique auquel il se rapporte. Les caractéristiques (localisation, couleur, texture, etc.) des régions recherchées sont déduites par l'analyse de ces germes. Généralement, ajouter ou supprimer des germes modifie le résultat produit, permettant à l'utilisateur d'obtenir la segmentation qu'il souhaite après quelques itérations.

Du point de vue des indications demandées à l'utilisateur, deux catégories de méthodes peuvent être distinguées, selon que l'algorithme cherche les contours présents dans l'image ou qu'il tente de trouver une partition en régions. Lorsqu'une méthode de segmentation interactive cherche les contours des éléments composant l'image, il est généralement demandé à l'utilisateur de donner certains points appartenant à ces contours. L'une des méthodes les plus populaires est sans doute les ciseaux intelligents

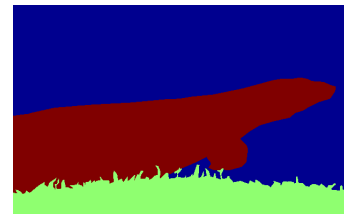
*Financement par ANR-11-LABX-0040-CIMI au sein du programme ANR-11-IDEX-0002-02.



(a) Image d'un reptile dans l'eau, mise à disposition par Santner [16].



(b) Segmentation dans l'optique d'une application de retouche d'image : l'élément principal (le reptile) est extrait du fond.



(c) Segmentation dans l'optique d'une application de compréhension d'image : chaque région correspond à un élément sémantique (reptile, eau et herbe).

FIGURE 1 – Plusieurs segmentations possibles pour une même image.

(ou lasso magnétique) [12] inclus depuis de nombreuses années dans des logiciels tels que Photoshop ou GIMP. Les méthodes recherchant les régions demandent, pour la plupart, de tracer des traits sur les différentes régions à extraire [4, 16, 17, 18]. Les pixels ainsi attribués à chaque région sont appelés germes.

Une fois les indications données par l'utilisateur, l'algorithme les utilise pour en déduire les caractéristiques des différentes classes qui doivent être trouvées par l'opération de segmentation. Quelques auteurs, à l'instar de Boykov *et al.* [4] ou Grady [7], utilisent simplement le niveau de gris des pixels ainsi que leur localisation pour segmenter l'image en plusieurs régions homogènes. D'autres algorithmes [3, 6, 8, 9, 14] utilisent la couleur du pixel, le plus souvent exprimée dans l'un des trois espaces colorimétriques suivants : RGB, HSV et Lab. Récemment, certains travaux [8, 16, 18] ont commencé à étudier l'utilisation d'une information de texture, avec des descripteurs tels que les textons, les tenseurs de structure ou les histogrammes de gradients orientés.

Un dernier aspect essentiel des méthodes de segmentation sémantique concerne la manière dont elles modélisent le problème de segmentation. Deux types d'approches se distinguent : les algorithmes par croissance de régions [13, 15, 18] et les algorithmes par minimisation d'une énergie [4, 7, 16].

Les algorithmes par croissance de régions reposent sur une première étape d'initialisation où les germes sont considérés comme les points de départ des régions recherchées. Ces embryons de régions sont ensuite agrandis en agglomérant les pixels voisins qui satisfont un critère de similarité. À chaque ajout d'un nouveau pixel, les caractéristiques des régions sont mises à jour. Ce processus est répété jusqu'à ce que tous les pixels appartiennent à une région.

Les algorithmes par minimisation d'énergie modélisent le problème de segmentation par une fonction dont la valeur dépend des caractéristiques de l'image et de la segmentation obtenue. Cette fonction atteint son minimum pour une segmentation optimale. Les germes donnés par l'utilisateur sont intégrés comme des contraintes. Une définition classique de la fonction de minimisation consiste à s'assurer que sa valeur décroît lorsque les régions de la segmentation sont homogènes, que les bordures sont régulières et que ces régions sont cohérentes vis-à-vis des germes donnés par l'utilisateur.

L'algorithme que nous proposons appartient à la catégorie des méthodes recherchant une partition de l'image en régions. Elle repose sur une première étape de regroupement des pixels en petites régions connexes, les superpixels, puis sur une classification de ces derniers. Il est à notre connaissance nouveau d'utiliser directement le résultat du classificateur. Utilisé pour classer des pixels ce dernier produit une segmentation bruitée qui nécessite d'être améliorée [16]. Dans le cas de SCIS, un post-traitement est superflu, l'homogénéité de la segmentation résultat étant assurée par le fait que les entités classées soient des superpixels.

3 Segmentation interactive par classification de superpixels

Dans cette section, nous décrivons la nouvelle méthode de segmentation interactive SCIS que nous avons implémentée comme greffon pour le logiciel de manipulation d'images GIMP¹. Les germes sont donnés à l'aide d'un pinceau. Les germes pour deux éléments sémantiques différents doivent avoir des couleurs différentes (par exemple bleu pour la classe ciel, vert pour la classe herbe, etc.) mais deux régions distinctes peuvent appartenir à la même classe même si leurs pixels ne sont pas connectés. Pour que SCIS fournisse un résultat correct, les germes doivent être répartis de manière uniforme dans l'image.

3.1 Algorithme général

Soit une image de M pixels où chaque pixel est associé à un vecteur contenant ses caractéristiques (niveau de gris, couleur, etc.). Cette image peut donc être décrite comme un ensemble P de M vecteurs, avec p_i le vecteur contenant les caractéristiques du $i^{\text{ème}}$ pixel. Tout d'abord, nous groupons ces pixels en petits ensembles homogènes, nommés superpixels. Pour chaque superpixel, nous extrayons ses caractéristiques.

La seconde partie de l'algorithme est itérative : aussi longtemps que l'utilisateur ajoute ou supprime des germes, la segmentation est mise à jour. Nous commençons par analyser les germes donnés par l'utilisateur afin d'en déduire K , le nombre de classes : comme les germes sont des traits colorés, nous avons une classe par couleur de germe. Nous créons un ensemble de $K + 1$ éti-

1. <http://www.gimp.org/>

quettes $\{0, \dots, K\}$, avec $1, \dots, K$ les étiquettes représentant chaque classe et 0, une étiquette vide. Ainsi nous pouvons assigner à chaque pixel une étiquette j telle que $j > 0$ si p est un germe et $j = 0$ sinon. Pour chaque classe, nous créons S_j , l'ensemble des superpixels incluant au moins un pixel avec l'étiquette j et dont tous les autres pixels ont soit l'étiquette j , soit l'étiquette 0. Ensuite, nous utilisons une méthode de classification multiclassée par apprentissage, entraînée pour chaque classe j avec les caractéristiques des superpixels de l'ensemble S_j . À la fin de cette étape, certains superpixels n'ont pas de classe, soit parce qu'ils ne contiennent aucun germe, soit parce qu'ils contiennent des germes appartenant à des classes différentes. Nous classons ces superpixels grâce au classifieur, qui fournit pour chaque superpixel la classe la plus probable. Si les probabilités pour plusieurs classes sont identiques, nous sélectionnons l'une des classes de manière aléatoire. Enfin, nous assignons aux pixels la classe attribuée au superpixel auquel ils appartiennent. Un résumé de notre méthode est donné dans l'algorithme 1.

Algorithme 1 Superpixel Classification-based Interactive Segmentation

ENTRÉES :
 Image /* image couleur */
 Germes /* localisations et étiquettes des germes */
 PROCÉDURE :
 Sur-segmenter l'image en superpixels et créer S , l'ensemble de tous les superpixels.
 Extraire les caractéristiques des superpixels
Tant que l'utilisateur donne ou supprime des germes
Faire
 Déduire K le nombre de classes en analysant les germes
 Pour chaque classe j **Faire**
 Créer l'ensemble S_j des superpixels ne contenant que des pixels avec l'étiquette j ou des pixels sans étiquette
 Fin Pour
 Entraîner le classifieur avec $S_{train} = \bigcup_{j=1, \dots, K} S_j$
Pour chaque superpixel $s \in S - S_{train}$ **Faire**
 Utiliser le classifieur pour prédire sa classe
Fin Pour
Pour chaque superpixel $s \in S$ **Faire**
 Pour chaque pixel $p \in s$ **Faire**
 Assigner la classe de s à p
 Fin Pour
Fin Pour
Fin Tant que

3.2 Sur-segmentation en superpixels

Les superpixels sont le résultat d'une sur-segmentation de l'image : les bordures de l'image doivent correspondre aux bordures des superpixels, mais un même élément peut être segmenté en plusieurs superpixels. Parce que les superpixels sont considérablement moins nombreux que les pixels, entraîner et utiliser une méthode de classification sur un ensemble de superpixels se révèle considé-

ablement plus rapide que le même traitement au niveau des pixels. Toutefois les erreurs dans la sur-segmentation (c'est-à-dire un superpixel chevauchant deux éléments que l'utilisateur souhaite séparer) ne peuvent pas être corrigées, quel que soit le nombre de germes ajoutés ou enlevés. Nos expérimentations menées à partir de la vérité terrain créée par McGuinness *et al.* [11] nous ont conduits à choisir l'algorithme de Felzenszwalb *et al.* [5], à la fois pour sa rapidité et pour son faible taux d'erreurs dans la sur-segmentation.

Chaque superpixel est ensuite décrit par sa couleur moyenne calculée dans l'espace colorimétrique RGB et la position de son barycentre, ce qui donne un vecteur à 5 dimensions. Chaque composante de ce vecteur est normalisée. L'évaluation que nous décrivons dans la section 4 montre que, malgré leur simplicité, ces caractéristiques se sont révélées suffisantes pour obtenir des résultats satisfaisants. Elles ont en outre l'avantage d'être rapidement calculées.

3.3 Méthode de classification

Les travaux de Santner *et al.* [16] ainsi que ceux de Xu *et al.* [17] ont montré que les séparateurs à vaste marge (SVM) et les forêts d'arbres décisionnels (RDT), sont deux types de méthodes de classification avec apprentissage qui peuvent être utilisées avec succès dans un algorithme de segmentation interactive. Nous les avons donc testées toutes les deux. Pour le SVM, nous avons utilisé une méthode de type C-SVM, implémentée dans la bibliothèque libSVM². Nous avons choisi d'utiliser un noyau gaussien. Pour la RDT, nous avons utilisé l'implémentation de la bibliothèque alglib³.

Dans un premier temps, nous avons comparé ces deux méthodes de classification à partir d'un sous-ensemble des images et des segmentations de référence fournies par Santner [16], produisant deux variantes de SCIS : SCIS-SVM et SCIS-RDT. Nous les avons testées avec des couples de paramètres différents et nous avons analysé l'évolution de leur comportement, lorsque ces paramètres sont modifiés. Nous avons sélectionné quatre critères : le temps moyen d'exécution, une mesure de précision (ACC), le pourcentage de germes et le temps total mis par l'utilisateur pour obtenir la segmentation désirée. La mesure de précision se base sur une segmentation de référence, produite par un être humain. Elle correspond à la proportion de résultats corrects par rapport au nombre total de résultats :

$$ACC = \frac{\sum_1^N T_j}{\sum_1^N T_j \cup F_j} \quad (1)$$

avec T_j le nombre de vrais positifs, F_j le nombre de faux positifs et N le nombre de classes dans la segmentation de référence.

Le pourcentage de germes doit être compris comme le pourcentage minimal de pixels qui doivent être sélectionnés par l'utilisateur pour obtenir la segmentation désirée.

2. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>

3. <http://www.alglib.net/>

Il est évidemment impossible d'être certain d'avoir obtenu ce pourcentage minimal, aussi nous nous sommes contentés de donner aussi peu de germes que possible, nous arrêtant dès que la segmentation semblait stabilisée. Le temps moyen mis par l'utilisateur pour donner ces germes correspond au temps écoulé entre l'instant où le premier germe a été donné et celui où le dernier germe a été donné, moins le temps d'exécution. Nos tests ont montré que les résultats de SCIS étaient optimaux en employant un SVM, de paramètres $C = 4$, $\gamma = 4$. Les différences entre les deux classificateurs se situent davantage au niveau de la qualité des résultats, les temps de calcul étant similaires.

4 Étude comparative

Nous avons comparé SCIS à quatre méthodes de segmentation interactive : les deux méthodes ayant obtenu les meilleurs résultats dans l'évaluation de McGuinness [11] (BPT et IGC) ainsi que celle d'Arbelaez [2] (CDHIS) et celle de Santner *et al.* [16] (TSRFTV).

4.1 Données de référence fournies par McGuinness *et al.*

Dans leur étude [11], McGuinness *et al.* comparent l'efficacité de quatre méthodes de segmentation interactive [1, 4, 6, 15] pour la résolution de problèmes à deux classes. Les 96 images qu'ils ont utilisées sont un sous-ensemble de la base de données de Berkeley pour les problèmes de segmentation [10], choisis pour être représentatif d'une large variété de problèmes de segmentation. Au total, 100 segmentations de référence ont été produites par des êtres humains. La qualité des binarisations, produites par les méthodes de segmentation interactive, est évaluée grâce à deux mesures : une mesure de précision des bordures et une mesure de précision des objets.

La mesure de précision des bordures est définie de la manière suivante :

$$ACC_{Boundary} = 100 \frac{\sum_x \min(\tilde{B}_G(x), \tilde{B}_M(x))}{\sum_x \max(\tilde{B}_G(x), \tilde{B}_M(x))} \quad (2)$$

avec B_G et B_M les contours internes des pixels respectivement dans la segmentation de référence et dans la segmentation résultat et \tilde{B}_G ainsi que \tilde{B}_M les mêmes ensembles mais étendus grâce à la théorie de la logique floue, comme cela est décrit dans l'article [11].

La mesure de précision des objets est définie de la manière suivante :

$$ACC_{Object} = 100 \frac{|G_O \cap M_O|}{|G_O \cup M_O|} \quad (3)$$

avec M_O l'ensemble des pixels attribués à la forme par l'algorithme, G_O l'ensemble des pixels appartenant à la forme dans la segmentation de référence et $|S|$ la cardinalité de l'ensemble S .

Nous avons comparé les résultats de SCIS avec les résultats obtenus par les deux méthodes obtenant les meilleurs résultats dans l'étude de McGuinness *et al.*, à savoir IGC et BPT, ainsi qu'avec la méthode CDHIS d'Arbelaez *et al.*.

TABLE 1 – Score moyen de précision des bordures et score moyen de précision des objets (équations 2 et 3).

Algorithme	$ACC_{Boundary}$	ACC_{Object}
BPT	78%	92%
IGC	77%	92%
CDHIS	70%	91%
SCIS	82%	94%

Nous avons évalué SCIS et CDHIS sur les mêmes images que McGuinness. Nous avons utilisé les mêmes segmentations de référence et les mêmes mesures. De plus, à l'instar de McGuinness *et al.*, nous avons restreint les utilisateurs à une durée de deux minutes par image. Enfin, nous avons utilisé les mêmes germes pour SCIS et CDHIS.

Le tableau 1 montre les résultats obtenus par chacune des méthodes : BPT [15], IGC [4], CDHIS [2] et SCIS. Ces scores correspondent à la précision moyenne des bordures et à la précision moyenne des objets, une fois que l'utilisateur a fini d'ajouter ou de retirer des germes ou que le temps maximal autorisé pour une image est atteint. Selon les deux critères utilisés par McGuinness *et al.*, SCIS arrive en tête. Les temps de calculs pour segmenter une image à partir d'un ensemble de germes sont de quelques millisecondes.

4.2 Données de référence de Santner

Dans leur article [16], Santner *et al.* proposent une nouvelle méthode de segmentation interactive pour des problèmes multiclassés. Ils mettent également à disposition le matériel nécessaire pour refaire leurs tests, à savoir les 243 images qu'ils ont utilisées (de 625×391 pixels) ainsi que les 262 segmentations de référence correspondant à ces images. Leur méthode est évaluée grâce à la mesure DICE :

$$DICE(I_{result}, I_{groundtruth}) = 100 \sum_{i=1}^N 2 \frac{|R_i \cap G_i|}{|R_i \cup G_i|} \quad (4)$$

où I_{result} correspond à la segmentation résultat, $I_{groundtruth}$ à la segmentation de référence, N au nombre de classes, R_i à l'ensemble des pixels de la $i^{\text{ème}}$ classe dans la segmentation résultat et G_i à l'ensemble des pixels de la $i^{\text{ème}}$ classe dans la segmentation de référence.

Santner *et al.* fournissent en outre les germes qu'ils ont utilisés pour évaluer leur méthode. Pour nos tests, nous avons produit deux ensembles supplémentaires de germes. Les germes du premier ensemble, $Seeds_{SCIS}^1$ ont été donnés de manière à ce que chaque élément sémantique ait des germes et que la répartition spatiale de l'ensemble des germes soit approximativement uniforme. L'utilisateur était autorisé à ajouter ou supprimer des germes sans contrainte de temps, Santner *et al.* n'en ayant pas imposée dans leur évaluation. Ces germes correspondent à ceux produits par un utilisateur, lors d'une utilisation classique de notre méthode.

Les germes du second ensemble, $Seeds_{SCIS}^2$, ont été produits en demandant à l'utilisateur de se contenter de don-

TABLE 2 – *Mesure DICE (équation 4) obtenue à partir des images et segmentations de référence de Santner. Un score élevé indique qu'un nombre important de pixels appartiennent aux mêmes classes dans les segmentations résultats et dans les segmentations de référence.*

Algorithme	Germes Santner	Seeds ¹ _{SCIS}	Seeds ² _{SCIS}
TSRFTV	93%	–	–
CDHIS	91%	91%	95%
SCIS	82%	98%	98%

ner une étiquette à certains pixels isolés dans l'image. Ce type d'interaction avec l'utilisateur se révèle extrêmement pénible : son seul intérêt est d'avoir produit des germes sensiblement moins nombreux que ceux de l'ensemble Seeds¹_{SCIS}.

Nous avons comparé les résultats obtenus par SCIS et CDHIS avec la variante de la méthode de Santner produisant les meilleurs scores (TSRFTV). Pour les méthodes SCIS et CDHIS nous avons calculé les scores obtenus avec les germes de Santner et avec nos deux ensembles de germes.

Le tableau 2 montre que les germes fournis par Santner, qui ne sont pas distribués de manière uniforme dans l'image, ne conviennent pas à SCIS. Les germes des ensembles Seeds¹_{SCIS} et Seeds²_{SCIS} améliorent considérablement les résultats obtenus par SCIS. Les germes de Seeds¹_{SCIS} ont deux différences essentielles avec ceux de Santner : leur répartition dans l'image est plus uniforme et ils sont largement plus nombreux. Au contraire, les germes de l'ensemble Seeds²_{SCIS} sont bien moins nombreux que ceux produits par Santner (en moyenne nos germes représentent 0.09% des pixels tandis que ceux de Santner représentent 0.37% des pixels). En revanche leur distribution spatiale est plus uniforme. Les résultats que nous avons obtenus sur ces deux ensembles assurent donc que l'augmentation des performances de SCIS provient de la localisation des germes et pas de leur nombre. Les méthodes de segmentation interactive requérant des temps d'exécution faibles afin d'être confortables à utiliser, nous en concluons que SCIS est une bonne alternative aux méthodes de Santner et d'Arbelaez, qui ont le défaut d'être assez coûteuses en temps de calcul et en utilisation de la mémoire, tandis que SCIS ne requiert que quelques millisecondes pour effectuer une itération à partir d'un ensemble de germes. Quant au temps nécessaire pour segmenter correctement une image, il dépend essentiellement de la complexité du problème posé, notamment du nombre d'éléments à séparer et de leur taille. Même avec des images comprenant une dizaine d'éléments trois à cinq minutes suffisent pour obtenir une segmentation satisfaisante.

5 Annotation de photographies du paysage

Adoptée en octobre 2000, la convention européenne du paysage reconnaît l'utilité sociale du paysage, en tant qu'élément essentiel à la qualité de vie. L'entrée en vi-

gueur de ce traité a donné lieu à de nombreuses initiatives telles que la mise en place d'observatoires photographiques du paysage. Concrètement, il s'agit d'analyser les mécanismes et facteurs de transformation d'un espace, à partir de fonds photographiques, généralement conçus par la re-photographie régulière d'un même lieu. L'étude de ces séries d'images par des géographes et des personnes en charge de l'aménagement du territoire peut être facilitée par des outils logiciels. Nous pensons que SCIS constitue un bon moyen pour identifier et localiser rapidement les éléments de paysage présents sur une photographie. Ces annotations ouvrent ensuite des perspectives intéressantes, notamment dans le cadre d'une analyse automatique des fonds photographiques afin de rechercher des dynamiques (reboisement, construction des bâtiments, etc.) ou des éléments particuliers (arbres, panneaux, etc.). La figure 2 donne un exemple d'annotations réalisées à partir des fonds de l'observatoire de l'autoroute A89⁴. Faute de l'existence d'une vérité terrain, nous n'avons à ce jour effectué que des tests qualitatifs, sur une dizaine de photographies.

6 Conclusion

Dans cet article, nous avons proposé une méthode de segmentation interactive multiclassée à la fois simple et efficace. Nous avons implémenté cet algorithme comme un greffon pour le logiciel de manipulation d'images GIMP, disponible à l'adresse suivante : <http://image.enfa.fr/scis/>. Nous avons démontré ses performances sur deux problèmes de référence : les données de référence de McGuinness et celles de Santner. Par ailleurs, des tests menés sur une dizaine de photographies, tendent à montrer que la méthode SCIS peut être employée avec succès pour annoter rapidement et précisément des photographies de paysage. Ces images pourront ensuite être utilisées pour constituer une base de données dans laquelle des éléments particuliers ou des dynamiques spécifiques seront recherchés.

Si les résultats obtenus par SCIS s'avèrent extrêmement encourageants, il n'en demeure pas moins que certaines améliorations peuvent être envisagées. Actuellement l'étape la plus coûteuse en temps de calcul et ayant le plus fort impact sur la qualité du résultat concerne la sur-segmentation de l'image. Nos recherches actuelles se poursuivent dans ce domaine.

Remerciements

Nous remercions Pedro Felzenszwalb de l'université Brown, pour l'implémentation de son algorithme⁵. Nous remercions Chih-Chung Chang et Chih-Jen Lin de l'université Nationale de Taiwan, pour leur bibliothèque pour l'apprentissage et la classification par SVM : libSVM⁶. Nous remercions « l'observatoire photographique des territoires du Massif Central » qui nous a permis d'utiliser ses photographies prises lors de la construction de l'autoroute A89. Le travail de thèse de Bérengère Mathieu est financé par le Laboratoire d'Excellence CIMI (Centre International de Mathématiques et d'Informatique).

4. <http://poptmc.free.fr/>

5. <http://cs.brown.edu/~pff/segment/>

6. <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>

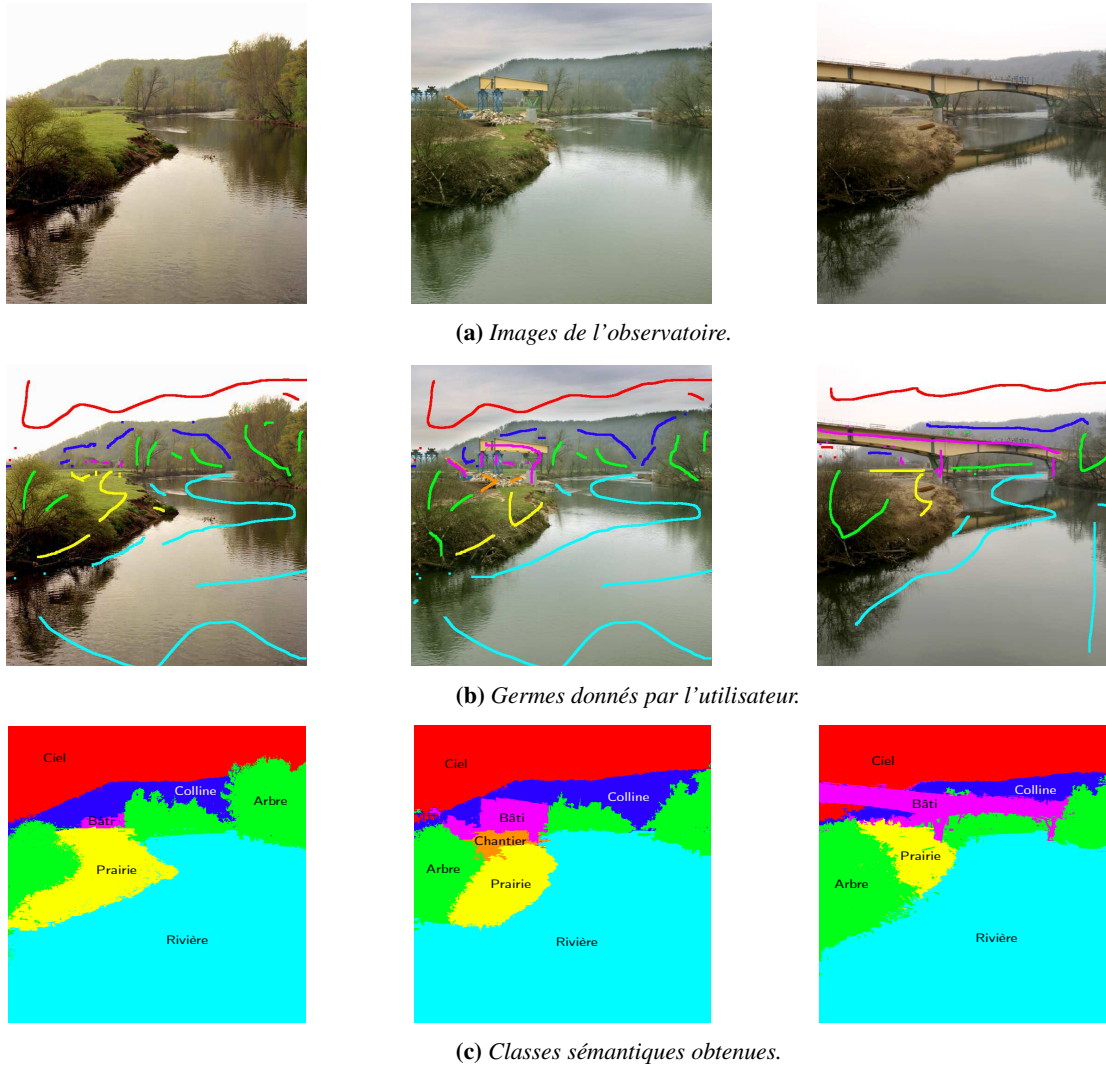


FIGURE 2 – Segmentation d'images issues d'un observatoire photographique du paysage.

Références

- [1] R. ADAMS et L. BISCHOF. Seeded region growing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(6):641–647, 1994.
- [2] P. ARBELAEZ, M. MAIRE, C. FOWLKES et J. MALIK. Contour detection and hierarchical image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(5):898–916, 2011.
- [3] A. BLAKE, C. ROTHER, M. BROWN, P. PEREZ et P. TORR. Interactive image segmentation using an adaptive gmmrf model. *In European Conference on Computer Vision (ECCV)*, pages 428–441, 2004.
- [4] Y. Y. BOYKOV et M.-P. JOLLY. Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. *In IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 105–112, 2001.
- [5] P. FELZENSZWALB et D. HUTTENLOCHER. Efficient graph-based image segmentation. *International Journal of Computer Vision*, 59(2):167–181, 2004.
- [6] G. FRIEDLAND, K. JANTZ et R. ROJAS. Siox: simple interactive object extraction in still images. *In IEEE International Symposium on Multimedia (ISM)*, pages 253–259, 2005.
- [7] L. GRADY. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.
- [8] Z. KUANG, D. SCHNIEDERS, H. ZHOU, K.-Y. K. WONG, Y. YU et B. PENG. Learning image-specific parameters for interactive segmentation. *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 590–597, 2012.
- [9] Y. LI, J. SUN, C. TANG et H.-Y. SHUM. Lazy snapping. *ACM Transactions on Graphics (ToG)*, 23(3):303–308, 2004.
- [10] D. MARTIN, C. FOWLKES, D. TAL et J. MALIK. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. *In IEEE International Conference on Computer Vision (ICCV)*, volume 2, pages 416–423, 2001.

- [11] K. MCGUINNESS et N. E. O'CONNOR. A comparative evaluation of interactive segmentation algorithms. *Pattern Recognition*, 43(2):434–444, 2010.
- [12] E. N. MORTENSEN et W. A. BARRETT. Intelligent scissors for image composition. *In International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, pages 191–198. ACM, 1995.
- [13] J. NING, L. ZHANG, D. ZHANG et C. WU. Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 43(2):445–456, 2010.
- [14] C. ROTHER, V. KOLMOGOROV et A. BLAKE. Grabcut: Interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics (TOG)*, 23(3):309–314, 2004.
- [15] P. SALEMBIER et L. GARRIDO. Binary partition tree as an efficient representation for image processing, segmentation, and information retrieval. *IEEE Transactions on Image Processing*, 9(4):561–576, 2000.
- [16] J. SANTNER, T. POCK et H. BISCHOF. Interactive multi-label segmentation. *In Asian Conference on Computer Vision (ACCV)*, volume 6492 de *Lecture Notes in Computer Science*, pages 397–410, 2011.
- [17] J. XU, X. CHEN et X. HUANG. Interactive image segmentation by semi-supervised learning ensemble. *In International Symposium on Knowledge Acquisition and Modeling (KAM)*, pages 645–648, 2008.
- [18] C. ZHOU et C. LIU. Interactive image segmentation based on region merging using hierarchical match mechanism. *In International Conference on Computer Science Service System (CSSS)*, pages 1781–1784, 2012.