

Segmentation de scènes urbaines par combinaison d'information

Marie-Anne Bauda^{1,2}

S. Chambon¹ M. Spangenberg² V. Charvillat¹

¹ IRIT - Université de Toulouse

² imajing sas

ORASIS juin 2013

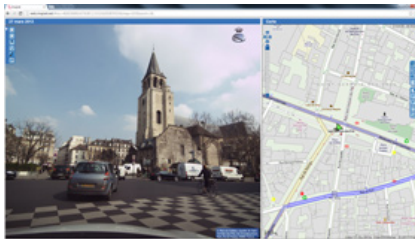


Contexte

- Thèse CIFRE (Avril 2012-2015)
- imajing propose une solution de relevé portable de réseaux de transports



imajbox ®



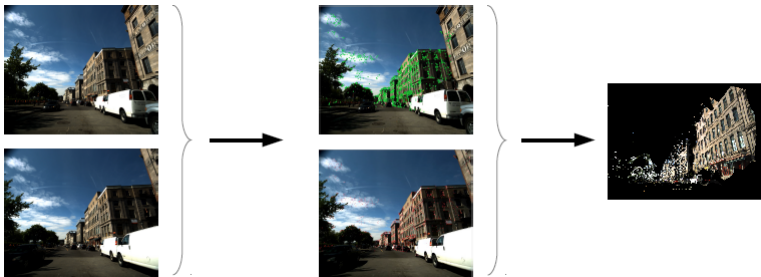
imajnet ®

Constat

Besoin de segmenter et de reconnaître automatiquement les entités présentes dans la scène

Données d'entrée

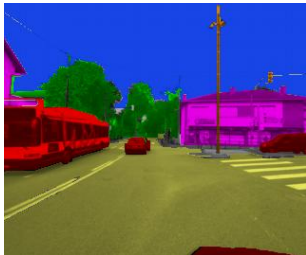
- Séquences d'images géo-référencées
- Données de navigation traitées
- Mise en correspondance de points d'intérêt



Les points d'intérêt détectés et appariés sur chaque paire d'images sont triangulés afin d'obtenir un nuage épars de points 3D.

Problématique

Segmentation de scènes extérieures par **combinaison d'information** 2D/3D en vue d'une description sémantique
➔ sur-segmentation en superpixels cohérents sur le triplet d'images pour une classification en trois classes (sol, ciel, objets verticaux)



Résultat espéré : vérité terrain



Contexte géométrique [Hoiem, 2005]

Sommaire

- 1 État de l'art
 - Segmentation
 - Superpixels et applications
- 2 Méthode proposée
 - Méthodologie
 - Algorithme
- 3 Résultats
- 4 Conclusions et perspectives

Sommaire

- 1 État de l'art
 - Segmentation
 - Superpixels et applications
- 2 Méthode proposée
- 3 Résultats
- 4 Conclusions et perspectives

En régions, en contours...

En régions :

→ on cherche des zones

- k-means
- mean-shift
[Comaniciu, 2002]
- regroupement de primitives
[Achanta, 2012]

En contours :

→ on cherche des discontinuités

- ligne de partage des eaux
[Vincent, 1991]
- carte de probabilité des contours [Chan, 2001]
- contours actifs [Chan, 2001]

Sur-segmentation en superpixels (SP)

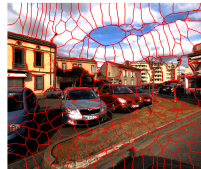
Historique : Le terme SP apparaît dans [Ren, 2007].

Définition : Un SP est une structure locale et cohérente représentant un objet ou une partie d'un objet.

- [Felzenszwalb, 2004]
- optimisation des variations intra et inter superpixels
- basée coupure de graphe normalisée



- [Moore, 2009]
- conservation de la régularité topologique
- carte de probabilité de contours + structure en bande



- [Mori, 2005]
- uniforme en taille et en surface
- basée coupure de graphe + détecteur de contours



- [Achanta, 2012]
- uniforme en taille et en surface
- regroupement de primitives

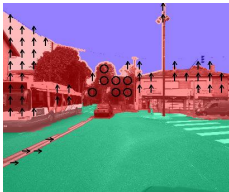
Résultats pour 200 superpixels.

Applications des SP

Compréhension de scène sur une image unique

[Hoiem, 2005]

- Application des SP de [Felzenszwalb, 2004] avec plusieurs hypothèses
- Classification avec un modèle d'apprentissage



Segmentation vidéo

[Grundmann, 2012]

- Application des SP de [Felzenszwalb, 2004]
- Utilisation de la cohérence spatio-temporelle



Sommaire

- 1 État de l'art
- 2 Méthode proposée
 - Méthodologie
 - Algorithme
- 3 Résultats
- 4 Conclusions et perspectives

Redondance d'information

- Utilisation de la redondance spatiale et temporelle
- Combinaison de l'information 2D/3D

→ triplet d'images

Type de résultats attendus : Niveau de l'objet, partie d'objet (relatif à la distance à la caméra)

Type de raisonnement : Superpixels (cohérence sur un support spatial)



Hypothèse

Scène plane par morceaux

Algorithme proposé

- 1 Sur-segmentation en SP sur l'image centrale du triplet
- 2 Classification « géométrique » des SP
- 3 Estimation des homographies existantes
- 4 Construction de nouveaux SP cohérents sur le triplet

(1) Sur-segmentation en superpixels : SLIC

SLIC, *Simple Linear Iterative Clustering* [Achanta, 2012]

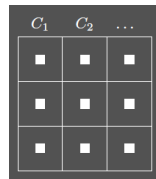
Algorithme

- 1 Initialisation : échantillonnage régulier de l'image
- 2 Évolution des centres de chaque superpixel
- 3 Renforcement de la connectivité des régions

(1.1) Échantillonnage régulier de l'image

Grille

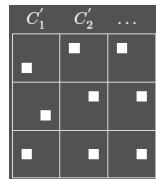
- K superpixels dans une image de N pixels
- Taille = $S \times S$ avec $S = \sqrt{\frac{N}{K}}$
- $C_k(L_k, a_k, b_k, x_k, y_k)$ centres



Ajustement de la grille

Pour chaque C_k

- Chercher le pixel C'_i de gradient le plus faible dans un voisinage $n \times n$
- $C_k \leftarrow C'_i$



Valeur de n ? $\rightarrow 3$

(1.2) Évolution des centres de chaque superpixel

Algorithme itératif

Répéter

- ❶ Calcul des superpixels
 - Pour chaque p_i , choisir C_k dans voisinage $2S \times 2S$ qui minimise

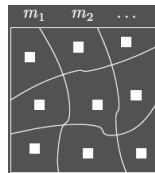
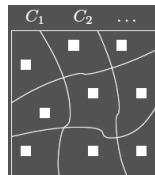
$$D_{\text{slic}} = d_{\text{lab}} + \frac{m}{S} d_{\text{xy}}$$

$d_{\text{lab}}, d_{\text{xy}}$: distances euclidiennes

m : poids entre la similarité de couleur et la proximité spatiale

- ❷ Mise à jour des centres (moyennes des attributs)
- ❸ Calcul de E (L_1 entre les anciens et nouveaux C_k)

Jusqu'à $E < \text{Seuil}$



(1) Sur-segmentation en superpixels : SLIC



Paramètres :

- $m = 20$
- $K = 20$

superpixels ➡ extraction locale de l'information 3D

(2) Classification « géométrique » des SP

Outil : Exploitation des points d'intérêt (PI) mis en correspondance (3D)

Soit $PI_i = \{PI \in SP_i\}$

- si $\#(PI_i) < k$ alors $\mathcal{C}(SP_i) = \text{information insuffisante}$
- sinon ($\#(PI_i) > k$) on cherche à estimer un plan dominant à l'aide de la méthode de RANSAC sur les données du nuage de points 3D
 - s'il existe un plan dominant (\vec{n}, d) alors $\mathcal{C}(SP_i) = \text{plan}$
 - sinon $\mathcal{C}(SP_i) = \text{non plan}$

Valeur de k ? $\rightarrow k > 4$

RANSAC : RANdom SAMple Consensus [Fischer, 1981]

(2) Classification « géométrique » des SP

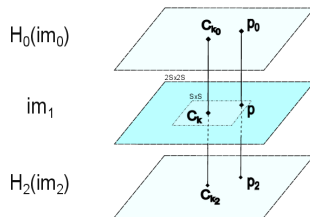


(3) Estimation des homographies

$$H_{0 \rightarrow 1} = K \left(R_{01} - \frac{t \vec{n}^\top}{d} \right) K^{-1}$$

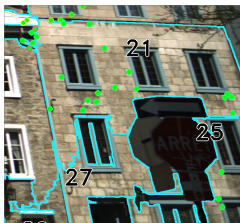
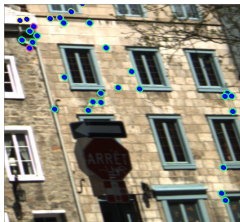
Pour chaque « SP plan » on estime les homographies existantes avec les images adjacentes correspondantes, H_0 avec l'image précédente et/ou H_2 avec l'image suivante :

- 1 superposition des données dans une image multi-canaux
- 2 calcul de nouveau SP de manière itérative et similaire à [Achanta, 2012]



On note $\text{diff}_0 = |im_1 - H_0(im_0)|$ et $\text{diff}_1 = |im_1 - H_2(im_2)|$

(3) Estimation des homographies



Différence homographique : diff_0

(4) Construction de nouveaux SP

$$D(C_k, p) = D_{slic}(C_k, p) + d_{plan}(p)$$

Si $\mathcal{C}(C_k) = plan$

$$d_{plan}(p) = \begin{cases} N_p & \text{si } \mathcal{C}(p) = plan \\ \max_{p' \in SP_k} N_{p'} & \text{sinon} \end{cases}$$

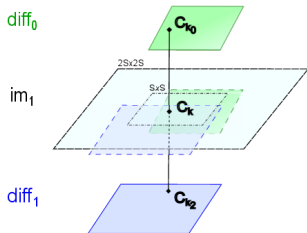
et

$$N_p = \text{diff}_0(p) \mathbb{1}_{H_0} + \text{diff}_1(p) \mathbb{1}_{H_2}(p)$$

Sinon

$$d_{plan}(p) = 0$$

$$\text{On note } \mathbb{1}_{H_i} = \begin{cases} 1 & \text{si } H_i \ni \\ 0 & \text{sinon} \end{cases}$$



Sommaire

- 1 État de l'art
- 2 Méthode proposée
- 3 Résultats**
- 4 Conclusions et perspectives

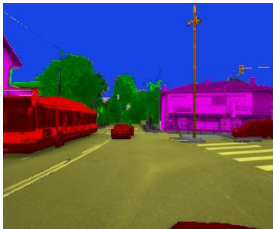
Corpus



Sources de variabilités
de la base de données :

- le type de réseau :
routes, chemins ferrés
- le type de trajectoire :
ligne droite, virage
- l'orientation de
l'imajbox
- la saisons et la
luminosité

Vérité terrain



Outil de segmentation manuelle
[Hoïem, 2005] :

- Précision au niveau du pixel (sur-segmentation possible)
- Description sémantique prédéfinie

Utilité :

- Qualifier nos résultats (erreur de sous-segmentation)
- Apprendre un modèle de classification

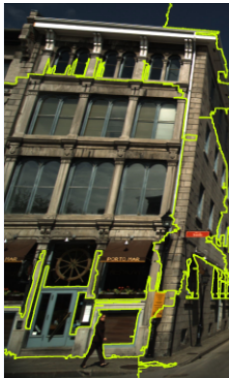
Évolue avec les données

Nos résultats vs. SLIC

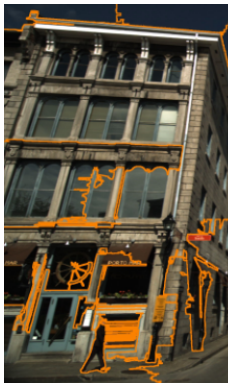


Comparaison entre les superpixels SLIC [Achanta, 2012] en orange et les nouveaux superpixels en jaune, en bleu les contours communs

Nos résultats vs. SLIC



Nouveaux superpixels



Superpixels SLIC

POUR

L'information prise en compte a une influence sur les nouveaux SP

CONTRE

Méthode constructive et destructive

Sommaire

- 1 État de l'art
- 2 Méthode proposée
- 3 Résultats
- 4 Conclusions et perspectives

Conclusions

- Méthode de sur-segmentation en SP prenant en compte un critère photométrique (D_{slic}) et **géométrique** (d_{plan})
- Résultats encourageants : la prise compte de la géométrie renforce la cohérence recherchée dans la scène
- Remise en cause de l'influence de l'étape 1 qui permet d'extraire localement l'information 3D pour estimer des zones planes

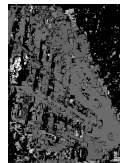
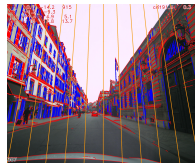
Perspectives

- ❶ Estimer la géométrie globale de la scène à l'aide d'un critère photométrique
- ❷ Utiliser l'information géométrique pour obtenir des SP cohérents géométriquement sur le triplet d'images

- Critères géométriques

- amélioration du RANSAC pour la détection de plan (estimation empirique des paramètres, intégration d'un critère photométrique)
- détection du « Up » et calcul de la ligne d'horizon

- Corrélation pour un nouveau critère de sur-segmentation



Merci pour votre attention.

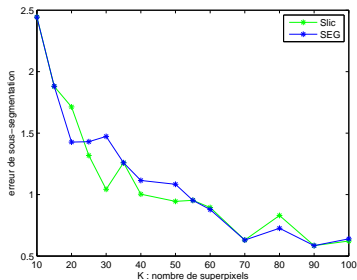
References I

- [Achanta, 2012] R. Achanta and A. Shaji and K. Smith and A. Lucchi and P. Fua and S. Susstrunk, SLIC Superpixels Compared to State-of-the-art Superpixel Methods, *IEEE Transactions on PAMI*, 2011.
- [Ayvaci, 2009] A. Ayvaci and S. Soatto, Motion segmentation with occlusions on the superpixel graph, *IEEE 12th ICCV Workshops*, 2009.
- [Bartoli, 2006] A. Bartoli, A Random Sampling Strategy for piecewise Planar Scene Segmentation, *CVIU*, 2006.
- [Chan, 2001] T. Chan and L. Vese, Active contours Without edges, *IEEE Transactions on Image Processing In Image Processing*, pp. 266-277, 2001.
- [Comaniciu, 2002] D. Comaniciu and P. Meer, Mean shift : a robust approach toward feature space analysis, *IEEE Transactions on PAMI*, pp. 603-619, 2002.
- [Felzenszwalb, 2004] P. Felzenszwalb and D. Huttenlocher, Efficient Graph-Based Image Segmentation, *IJCV*, pp. 167-181, 2004.
- [Fischer, 1981] M. Fischer and R. Bolles, Random Sample Consensus : A paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communication of the ACM*, pp. 381-397, 1981.
- [Fouhey, 2010] D. Fouhey and D. Scharstein and A. Briggs, Multiple Plane Detection in Image Pairs using J-Linkage, *IEEE ICPR*, 2010.
- [Grundmann, 2012] M. Grundmann and V. Kwatra, Efficient Hierarchical Graph Based Video Segmentation, *IEEE CVPR*, 2010.
- [Hoiem, 2005] D. Hoiem, Geometric Context from a Single Image, *IEEE ICCV*, pp. 654-661, 2005.
- [Kass, 1988] M. Kass and A. Witkin and D. Terzopoulos, Snakes : Active contour models, *IJCV*, pp. 321-331, 1988.
- [Martin, 2004] D. Martin and C. Fowlkes, Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues, *IEEE PAMI*, pp. 530-549, 2004.

References II

- [Mori, 2005] G. Mori, Guiding model search using segmentation, *IEEE ICCV*, pp. 1417–1423, 2005.
- [Moore, 2009] A. Moore and S. Prince, Scene shape priors for superpixel segmentation, *IEEE ICCV*, pp. 771-778, 2009.
- [Vincent, 1991] L. Vincent and P. Soille, Watersheds in Digital Spaces : An Efficient Algorithm Based on Immersion Simulations, *IEEE Transactions on PAMI*, 1991.
- [Ren, 2007] X. Ren and J. Malik, Tracking as Repeated Figure/Ground Segmentation, *IEEE CVPR*, 2007.
- [Wang, 2011] S. Wang, Superpixel tracking, *IEEE ICCV*, pp. 1323-1330, 2011.

Nos résultats vs. SLIC



Erreur de sous-segmentation pour SLIC (vert) et pour les résultats obtenus avec la méthode proposée (bleu)



SLIC [Achanta, 2012] en orange, les nouveaux superpixels en jaune et les contours communs en bleu