

Compréhension de scènes urbaines par combinaison d'information 2D/3D

Marie-Anne Bauda

Imajing
Vortex - IRIT - N7 - Université de Toulouse

Encadrants : Sylvie Chambon et Pierre Gurdjos
Directeur de thèse : Vincent Charvillat

13 juin 2016

imajing

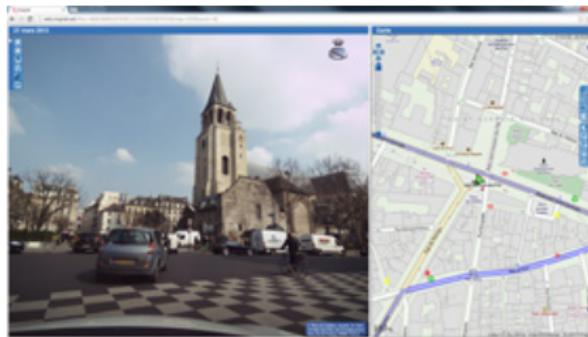


Contexte

- Thèse CIFRE
- Imajing : solution de relevé portable de réseaux de transports



imajbox®



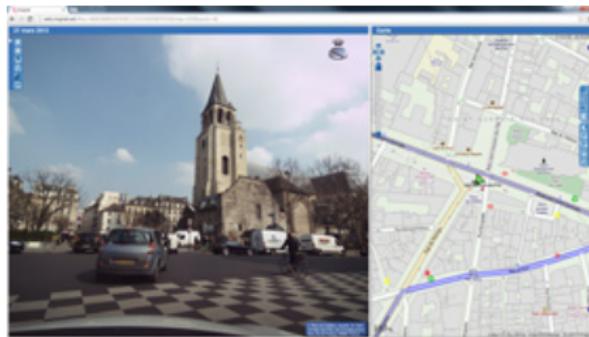
imajnet®

Contexte

- Thèse CIFRE
- Imajing : solution de relevé portable de réseaux de transports



imajbox ®



imajnet ®

Motivation

Besoin d'une segmentation sémantique en vue de reconnaître automatiquement des entités présentes dans la scène

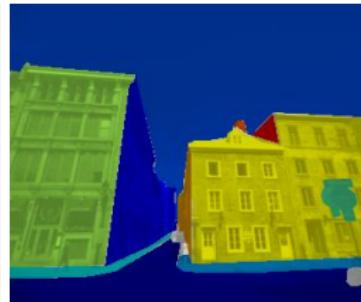
Problème

Segmentation de scènes urbaines par **combinaison d'information**
2D/3D en vue d'une description sémantique

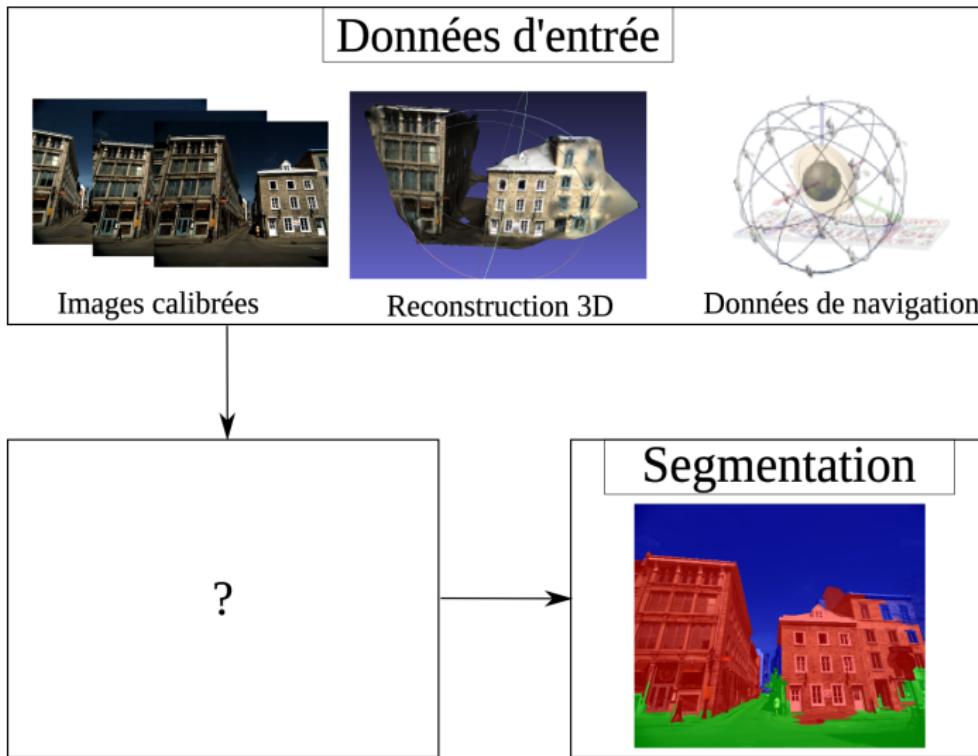


Problème

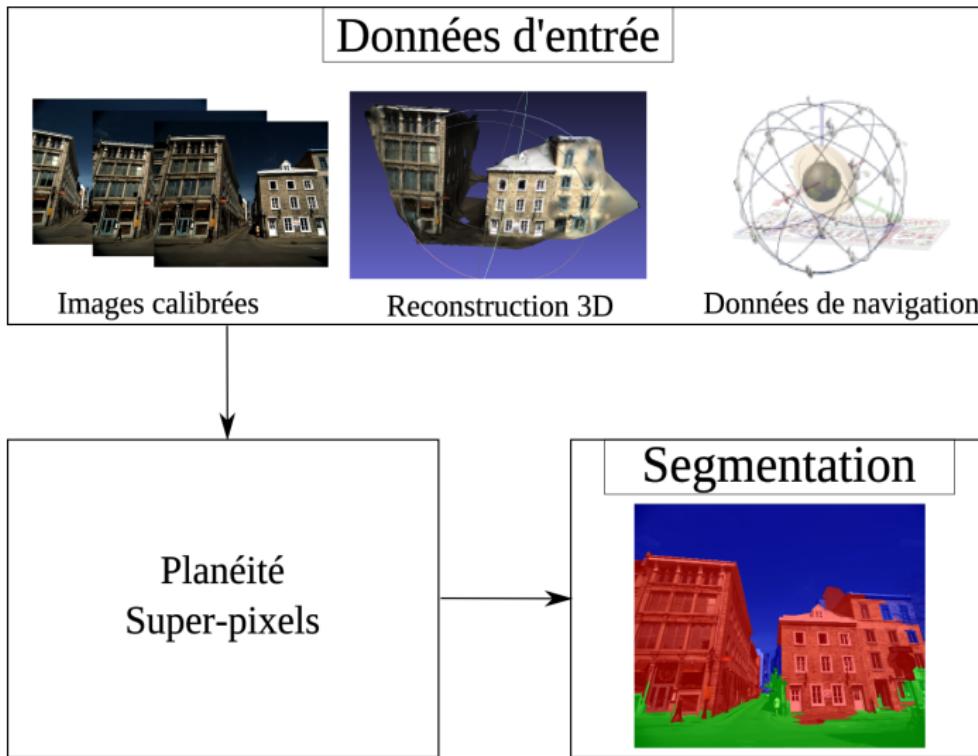
Segmentation de scènes urbaines par **combinaison d'information**
2D/3D en vue d'une description sémantique



Problème



Problème



Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

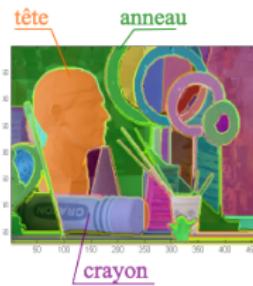
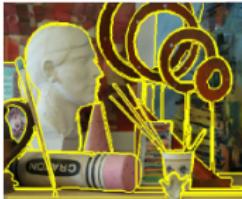
Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

Approches existantes de compréhension de scènes

Niveau de compréhension de la scène

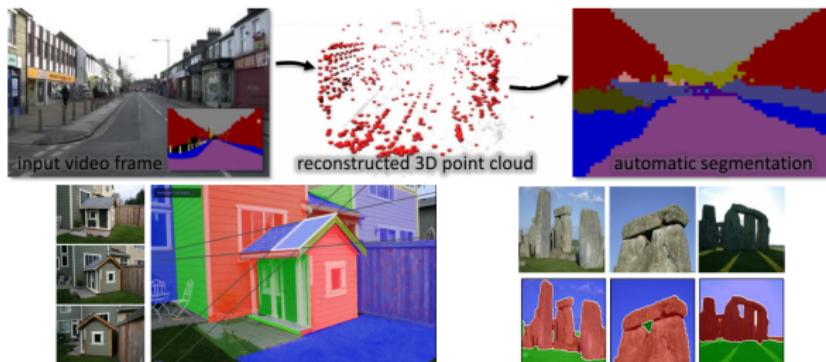
Pixel, super-pixel, objet, sémantique, etc.



Approches existantes de compréhension de scènes

Informations géométriques

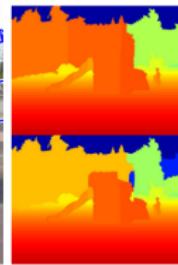
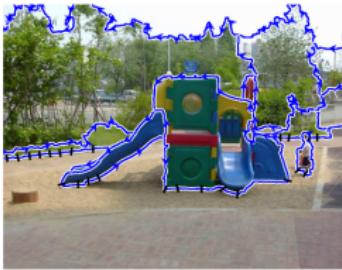
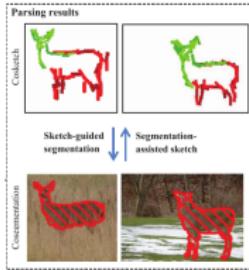
- Reconstruction 3D partielle, nuage de points ou carte de profondeur [Brostow, 2008]
- Planéité [Sinha, 2008]
- Co-segmentation [Joulin, 2012]



Approches existantes de compréhension de scènes

Informations sémantiques

- Silhouette et modèle de forme [Dai, 2013]
 - Position relative des objets [Hoiem, 2011]



Approches existantes de compréhension de scènes

Approches		Apprentissage	Planéité	Super-pixels	Multi-vues	Mono-vue
[Tighe, 2013]		x		x		
[Delage, 2007]		x			x	
[Alvarez, 2008]		x				x
[Hoiem, 2005]		x		x		x
[Fouhey, 2013]		x		x	x	x
[Grundmann, 2010]			x	x		
[Sinha, 2014]			x		x	
[Badrinarayanan, 2015]			x			x
[Bódis-Szomorú, 2014]			x	x	x	
[Raza, 2013]			x	x		x
[Gallup, 2008]			x		x	x



Approches existantes de compréhension de scènes

Approches	Apprentissage	Planéité	Super-pixels	Multi-vues	Mono-vue
[Tighe, 2013]	x		x		
[Delage, 2007]	x			x	
[Alvarez, 2008]	x				x
→ [Hoiem, 2005]	x		x		x
[Fouhey, 2013]	x		x	x	x
[Grundmann, 2010]		x	x		
[Sinha, 2014]	x			x	
[Badrinarayanan, 2015]	x				x
→ [Bódis-Szomorú, 2014]	x	x	x	x	
[Raza, 2013]	x	x			x
[Gallup, 2008]	x		x	x	x

Approche mono-vue [Hoiem, 2005]

Objectifs : Contexte géométrique "pop-up"

- Segmenter en trois classes : sol, ciel, objets verticaux
 - Rechercher l'orientation des surfaces au niveau du super-pixel
 - Couper et plier

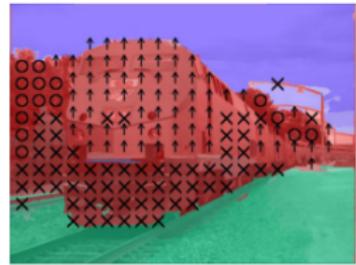
Approche mono-vue [Hoiem, 2005]

Éléments de l'approche

Approche probabiliste supervisée : vraisemblance de l'étiquette d'un super-pixel

- Sur-segmentation en super-pixels
 - Caractéristiques : des contours, régions (couleurs, textures), information 3D (lignes et jonctions)

Approche mono-vue [Hoiem, 2005]



Approche multi-vues [Bódis-Szomorú, 2014]

Objectifs : Approximation d'un modèle dense

Densification de la reconstruction du nuage épars

- Ajuster les paramètres de chaque plan
- Segmenter la scène en mettant en relation les points 3D et les régions
- Déterminer la visibilité des régions dans chaque image

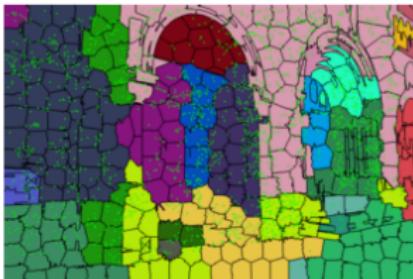
Approche multi-vues [Bódis-Szomorú, 2014]

Éléments de l'approche

Optimisation d'une fonction d'énergie par coupures de graphe

- Enveloppe convexe des super-pixels plans par morceaux
- Similarité de couleur et gradient de l'image (inter-vues et intra-vue)

Approche multi-vues [Bódis-Szomorú, 2014]



Formulation du problème

- Segmentation multi-vues en trois classes (sol, ciel, **objets verticaux**)
- Diversité de l'information disponible :
géométrique et photométrique
- Hypothèse de scène **plane/non-plane** par morceaux

Formulation du problème

- Segmentation multi-vues en trois classes (sol, ciel, **objets verticaux**)
- Diversité de l'information disponible : **géométrique et photométrique**
- Hypothèse de scène **plane/non-plane** par morceaux

Idée principale

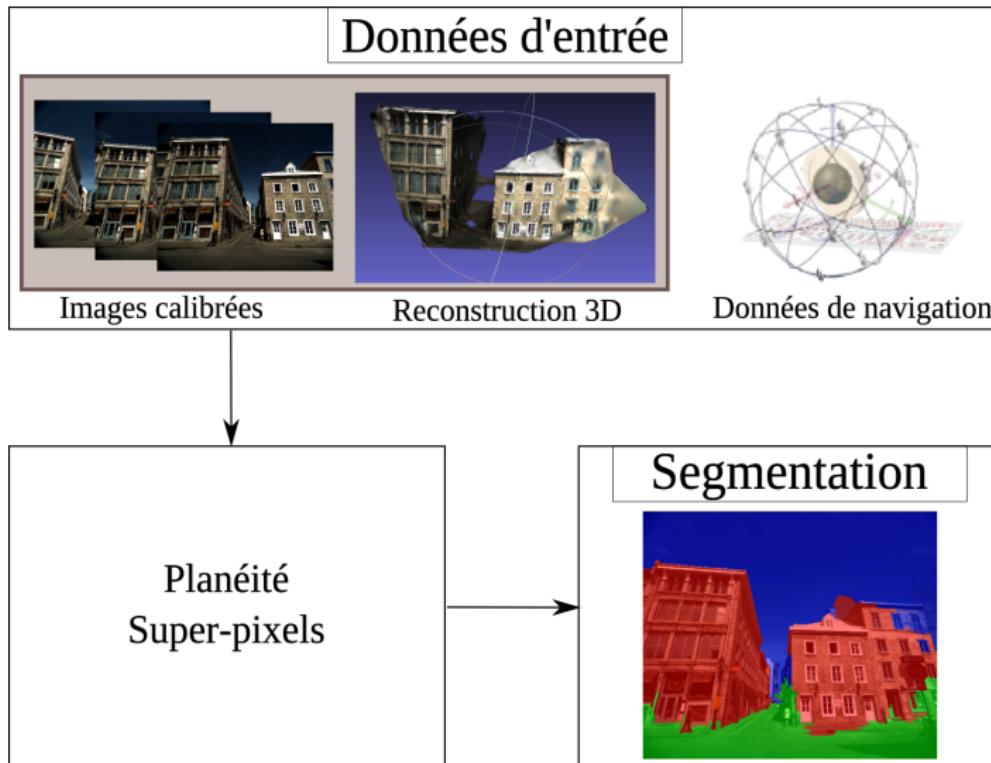
La géométrie de la scène peut aider la segmentation

→ sur-segmentation en **super-pixels** cohérents avec la géométrie de la scène pour une **segmentation en plans**

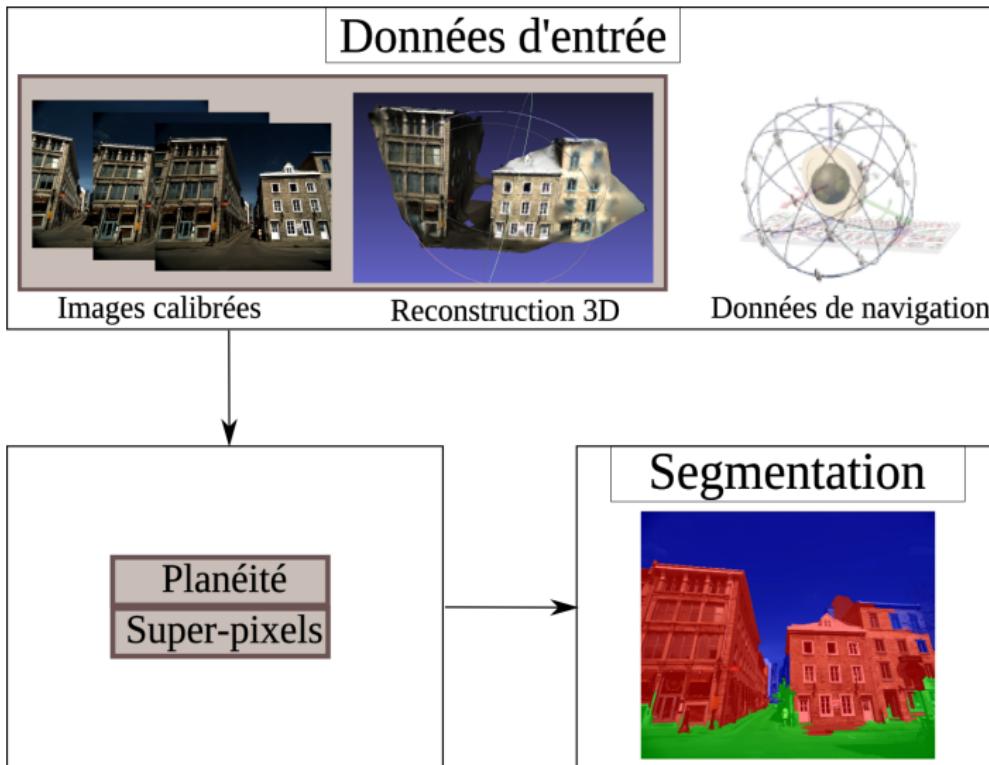
Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

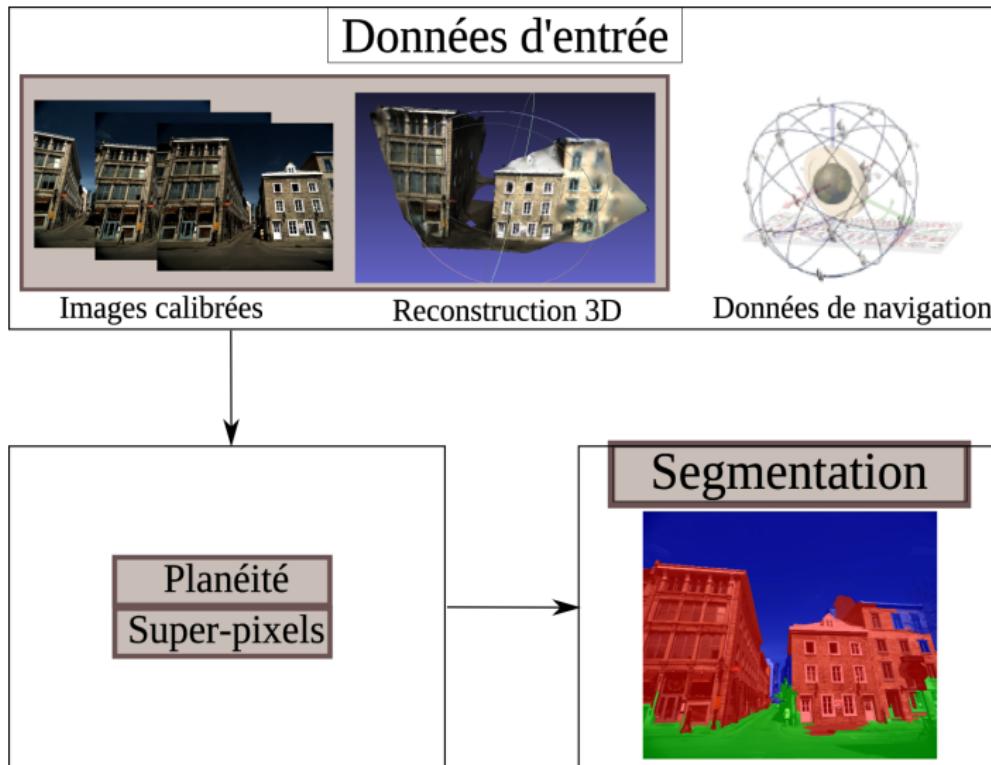
Vue d'ensemble



Vue d'ensemble



Vue d'ensemble



Plan et homographie

Plan Π

Trois points 3D définissent un plan Π de l'espace

Plan et homographie

Plan Π

Trois points 3D définissent un plan Π de l'espace

Homographie H_Π

$$H_\Pi \sim K^d (R - \frac{1}{d_\Pi} T n_\Pi^\top) (K^g)^{-1} \quad (1)$$

$(n_\Pi^\top, d_\Pi)^\top$, le vecteur homogène du plan Π

$K[R|T]$, les paramètres de la caméra

Plan et homographie

Plan Π

Trois points 3D définissent un plan Π de l'espace

Homographie H_Π

$$H_\Pi \sim K^d \left(R - \frac{1}{d_\Pi} T \mathbf{n}_\Pi^\top \right) (K^g)^{-1} \quad (1)$$

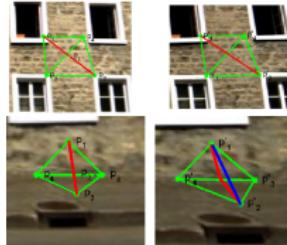
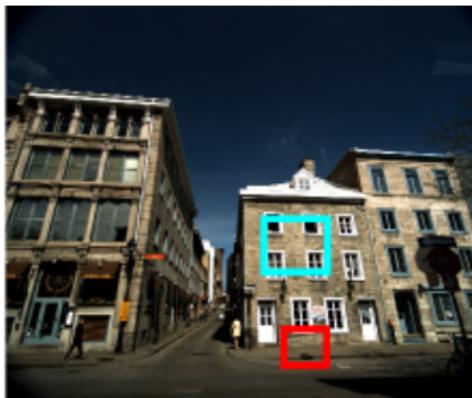
$(\mathbf{n}_\Pi^\top, d_\Pi)^\top$, le vecteur homogène du plan Π

$K[R|T]$, les paramètres de la caméra

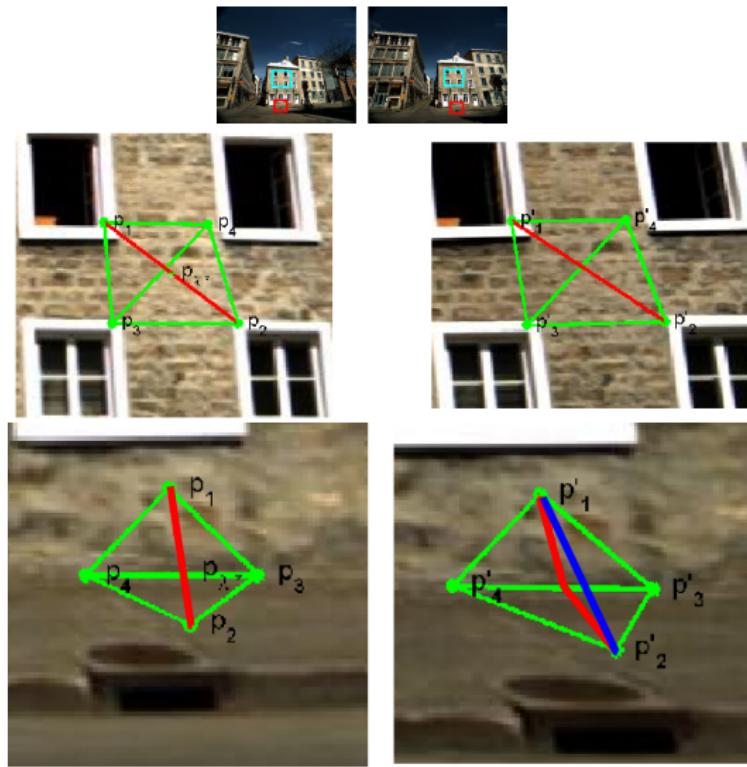
Vue recalée $\tilde{\mathbf{z}} = H_\Pi(\mathbf{z}')$

Image adjacente recalée par l'homographie induite par le plan de support

Cohérence photométrique ?



Cohérence photométrique ?



Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

Mesures de cohérence photométrique

Mesures de cohérence photométrique inter-images : IQA (*Image Quality Measure*)

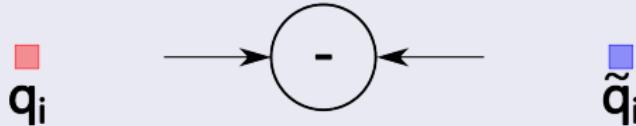
Analyse de deux types de mesures :

- Mesures s'appuyant sur la distance euclidienne
 - Mesures s'appuyant sur le cosinus de l'angle

Mesures s'appuyant sur la distance euclidienne

Erreur quadratique moyenne (Mean Square Error - MSE)

MSE

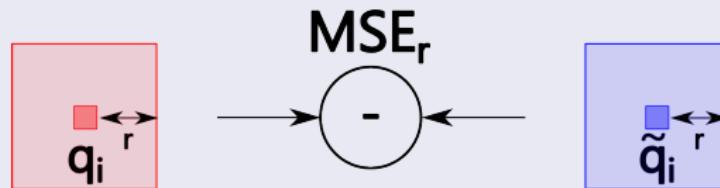


$$\text{MSE}(\mathbf{z}, \tilde{\mathbf{z}}) = \frac{1}{N} \sum_i (v_i - \tilde{v}_i)^2$$

- N , nombre de pixels dans \mathbf{z} et $\tilde{\mathbf{z}}$
- v_i , couleur de q_i dans l'image de référence
- \tilde{v}_i , couleur de \tilde{q}_i dans l'image recalée
- r , taille du voisinage considéré

Mesures s'appuyant sur la distance euclidienne

MSE généralisée

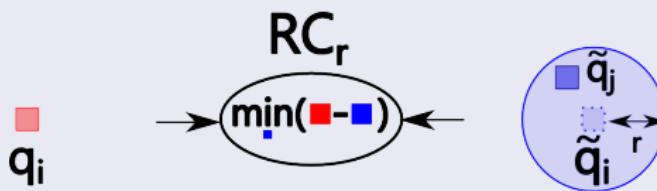


$$\text{MSE}_r(\mathbf{z}, \tilde{\mathbf{z}}) = \frac{1}{N} \sum_i \left[\frac{1}{(2r+1)^2} \sum_{j / |q_i - q_j| \leq r} (v_j - \tilde{v}_j)^2 \right]$$

- N , nombre de pixels dans \mathbf{z} et $\tilde{\mathbf{z}}$
- v_i , couleur de q_i dans l'image de référence
- \tilde{v}_i , couleur de \tilde{q}_i dans l'image recalée
- r , taille du voisinage considéré

Mesures s'appuyant sur la distance euclidienne

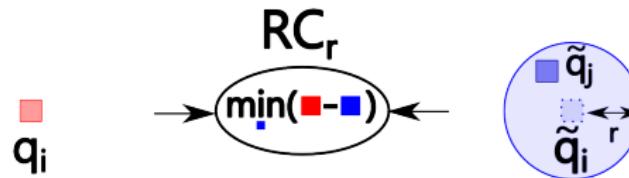
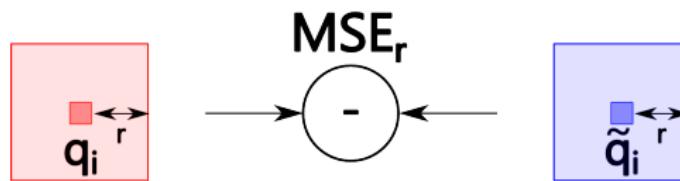
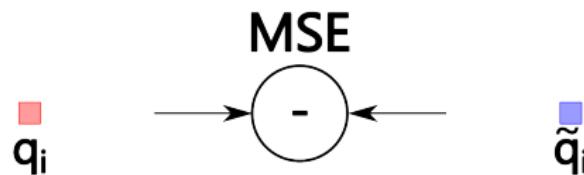
R-cohérence [Kutulakos, 2000]



$$RC_r(\mathbf{z}, \tilde{\mathbf{z}}) = \frac{1}{N} \sum_i \left(\min_{j / (q_i - q_j)^2 < r^2} |v_i - \tilde{v}_j| \right)^2$$

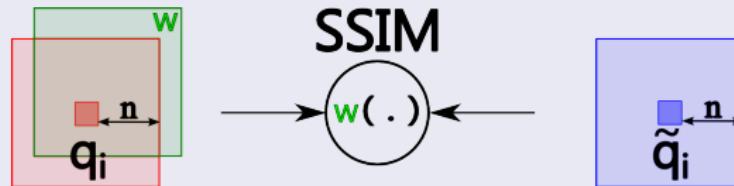
- N , nombre de pixels dans \mathbf{z} et $\tilde{\mathbf{z}}$
- v_i , couleur de q_i dans l'image de référence
- \tilde{v}_i , couleur de \tilde{q}_i dans l'image recalée
- r , taille du voisinage considéré

Mesures s'appuyant sur la distance euclidienne



Mesures s'appuyant sur le produit scalaire

Structural Similarity Measure [Wang, 2004]

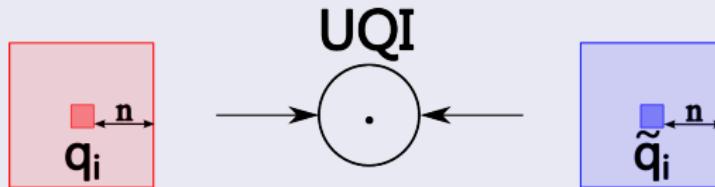


$$\begin{aligned}
 \text{SSIM}(\mathbf{z}, \tilde{\mathbf{z}}) &= l(\mathbf{z}, \tilde{\mathbf{z}}) c(\mathbf{z}, \tilde{\mathbf{z}}) s(\mathbf{z}, \tilde{\mathbf{z}}) \\
 &= \frac{2\mu_{\mathbf{z}}\mu_{\tilde{\mathbf{z}}} + \alpha}{\mu_{\mathbf{z}}^2 + \mu_{\tilde{\mathbf{z}}}^2 + \alpha} \frac{2\sigma_{\mathbf{z}}\sigma_{\tilde{\mathbf{z}}} + \beta}{\sigma_{\mathbf{z}}^2 + \sigma_{\tilde{\mathbf{z}}}^2 + \beta} \frac{\sigma_{\mathbf{z}\tilde{\mathbf{z}}} + \gamma}{\sigma_{\mathbf{z}}\sigma_{\tilde{\mathbf{z}}} + \gamma}
 \end{aligned}$$

- $\mu_{\mathbf{z}}$ (resp. $\mu_{\tilde{\mathbf{z}}}$), moyenne des v_i (resp. \tilde{v}_i) sur \mathbf{z} (resp. $\tilde{\mathbf{z}}$)
- $\sigma_{\mathbf{z}}$ (resp. $\sigma_{\tilde{\mathbf{z}}}$), écart-type de v_i (resp. \tilde{v}_i)
- $\sigma_{\mathbf{z}\tilde{\mathbf{z}}}$, covariance entre \mathbf{z} et $\tilde{\mathbf{z}}$

Mesures s'appuyant sur le produit scalaire

Universal Quality Image [Wang, 2002]

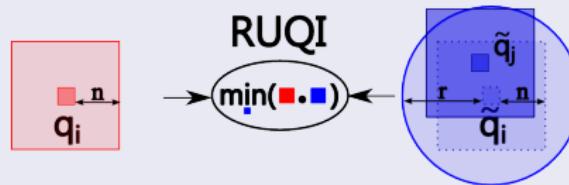


$$UQI(\mathbf{z}, \tilde{\mathbf{z}}) = \frac{4\sigma_{\mathbf{z}\tilde{\mathbf{z}}} \mu_{\mathbf{z}}\mu_{\tilde{\mathbf{z}}}}{(\sigma_{\mathbf{z}}^2 + \sigma_{\tilde{\mathbf{z}}}^2) [\mu_{\mathbf{z}}^2 + \mu_{\tilde{\mathbf{z}}}^2]}$$

- $\mu_{\mathbf{z}}$ (resp. $\mu_{\tilde{\mathbf{z}}}$), moyenne des v_i (resp. \tilde{v}_i) sur \mathbf{z} (resp. $\tilde{\mathbf{z}}$)
- $\sigma_{\mathbf{z}}$ (resp. $\sigma_{\tilde{\mathbf{z}}}$), écart-type de v_i (resp. \tilde{v}_i)
- $\sigma_{\mathbf{z}\tilde{\mathbf{z}}}$, covariance entre \mathbf{z} et $\tilde{\mathbf{z}}$

Mesures s'appuyant sur le produit scalaire

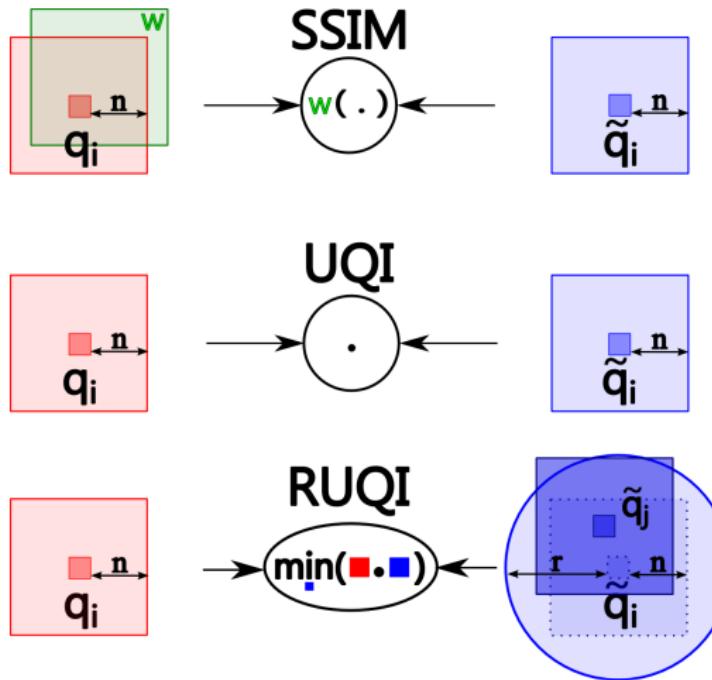
RUQI



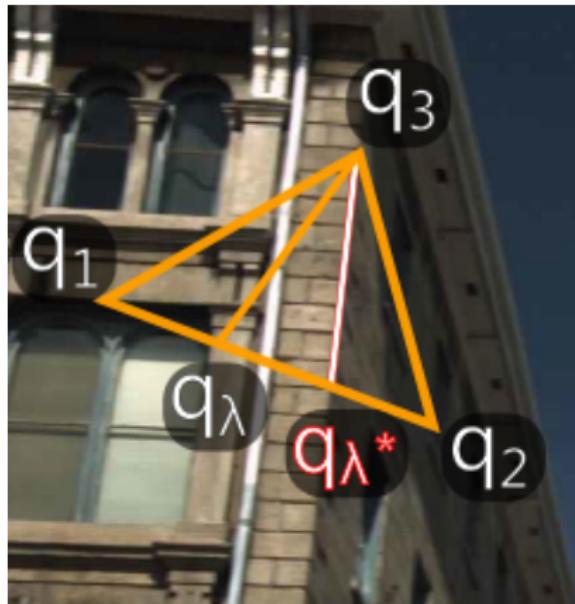
$$\text{RUQI}(\mathbf{z}, \tilde{\mathbf{z}}) = \frac{1}{N} \sum_i \left(\max_{j / (q_i - q_j)^2 < r^2} (\text{UQI}(\xi_i, \tilde{\xi}_j)) \right)$$

- $\mu_{\mathbf{z}}$ (resp. $\mu_{\tilde{\mathbf{z}}}$), moyenne des v_i (resp. \tilde{v}_i) sur \mathbf{z} (resp. $\tilde{\mathbf{z}}$)
- $\sigma_{\mathbf{z}}$ (resp. $\sigma_{\tilde{\mathbf{z}}}$), écart-type de v_i (resp. \tilde{v}_i)
- $\sigma_{\mathbf{z}\tilde{\mathbf{z}}}$, covariance entre \mathbf{z} et $\tilde{\mathbf{z}}$

Mesures s'appuyant sur le produit scalaire



Protocole d'évaluation des IQA

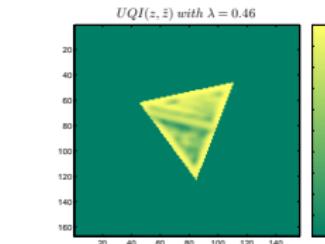
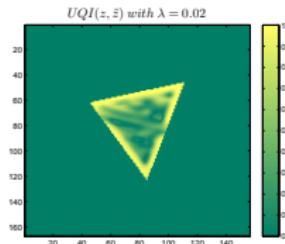
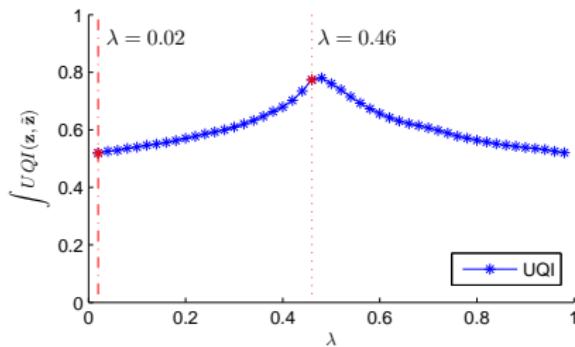
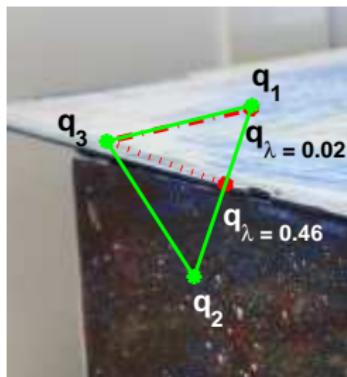


Protocole d'évaluation des IQA

→ *animation1_BD1 – 25_a93_UQI.gif*

Protocole d'évaluation des IQA

→ *animation1_BD1 – 25_a93_UQI.gif*



Protocole d'évaluation des IQA

Algorithme

Données : Images \mathbf{z} et \mathbf{z}' , $q_i \leftrightarrow q'_i, \forall i \in [1, \dots, 4]$

Résultat : Classification de \mathbf{z} en régions P/NP

Estimation de λ^* (valeur réelle)

Estimation des homographies

pour $\lambda \in [0, 1]$ **faire**

 Estimation de l'image recalée $\tilde{\mathbf{z}}$

 Calcul d'IQA($\mathbf{z}, \tilde{\mathbf{z}}$)

fin

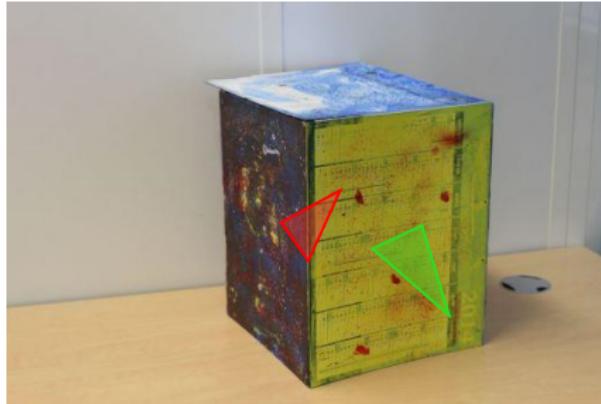
Classification en régions P/NP grâce à l'IQA

Protocole d'évaluation des IQA

Expérimentation

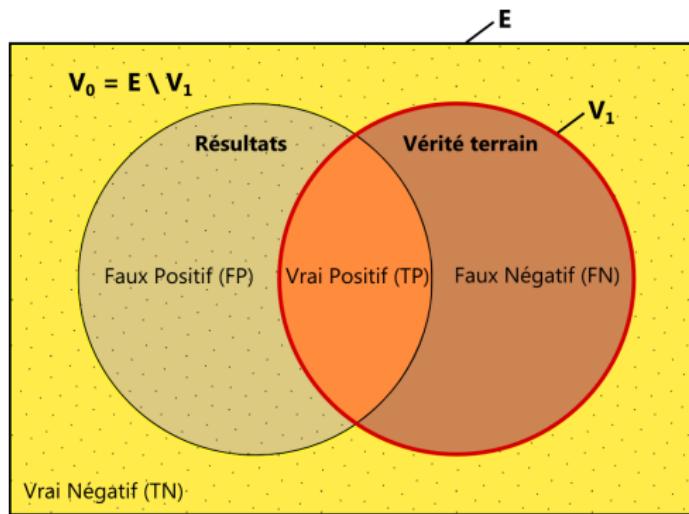
Données d'entrée - Données réelles avec et sans contrôle de la luminosité

Tests - 6 mesures sur 87 triangles (dont 58% NP)



Évaluation d'une classification

Définitions :



- vrais positifs (TP) pixels étiquetés à raison dans la classe d'intérêt
- vrais négatifs (TN) pixels étiquetés à raison dans l'autre classe
- faux positifs (FP) pixels étiquetés à tort dans la classe d'intérêt
- faux négatifs (FN) pixels étiquetés à tort dans l'autre classe

Évaluation d'une classification

Critères d'évaluation :

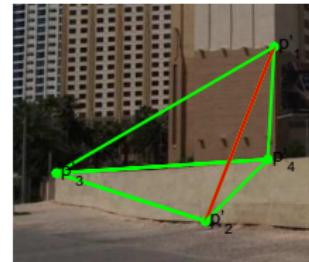
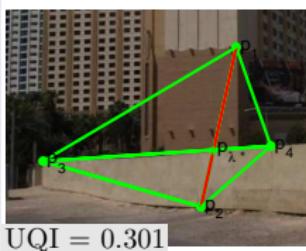
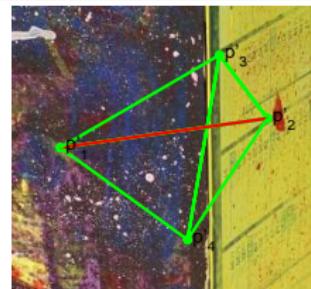
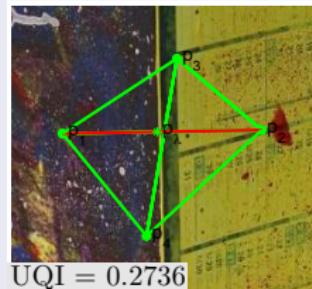
$$\text{Précision } Pr = \frac{TP}{TP + FP} \quad (2)$$

$$\text{Rappel } R = \frac{TP}{TP + FN} \quad (3)$$

$$\text{Taux de faux positifs } FPR = \frac{FP}{FP + TN} \quad (4)$$

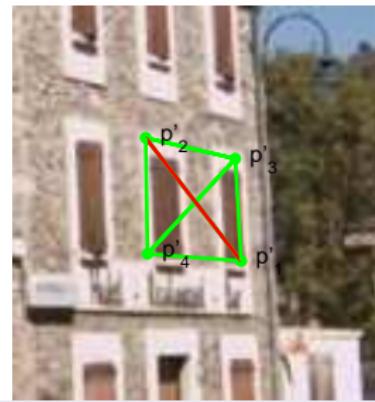
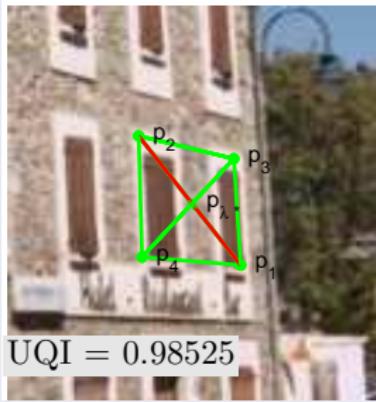
Résultats

Zones non-planes

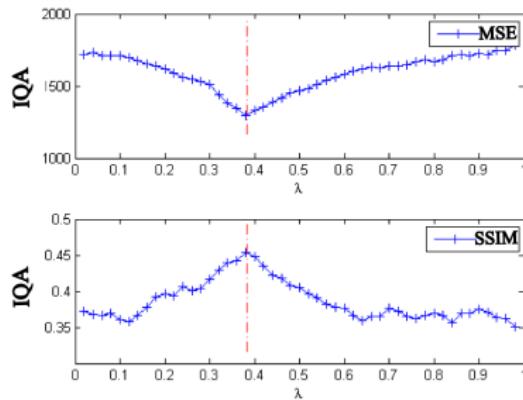
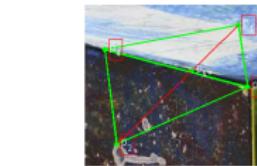
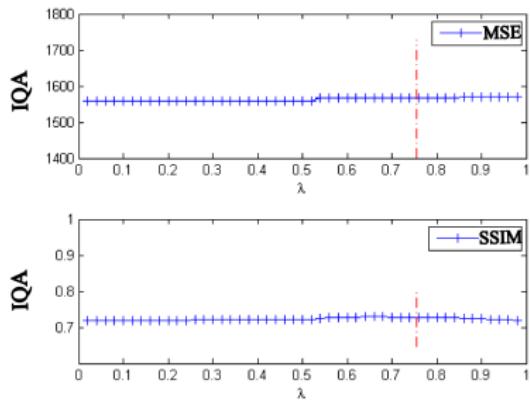
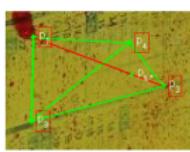


Résultats

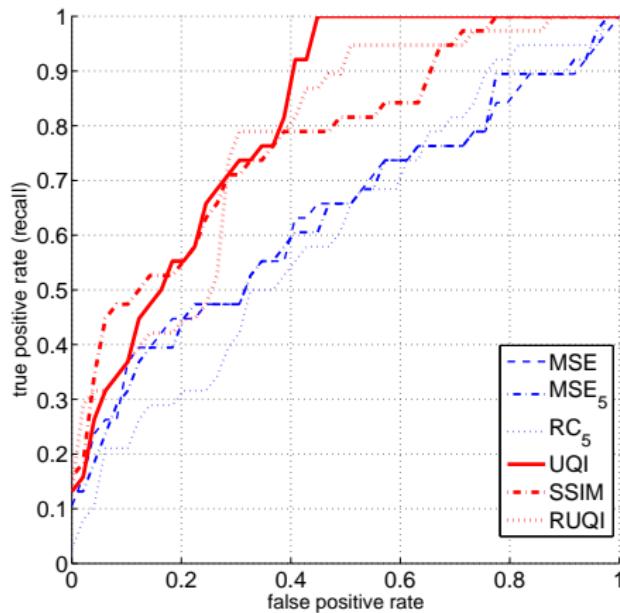
Zones planes



Résultats

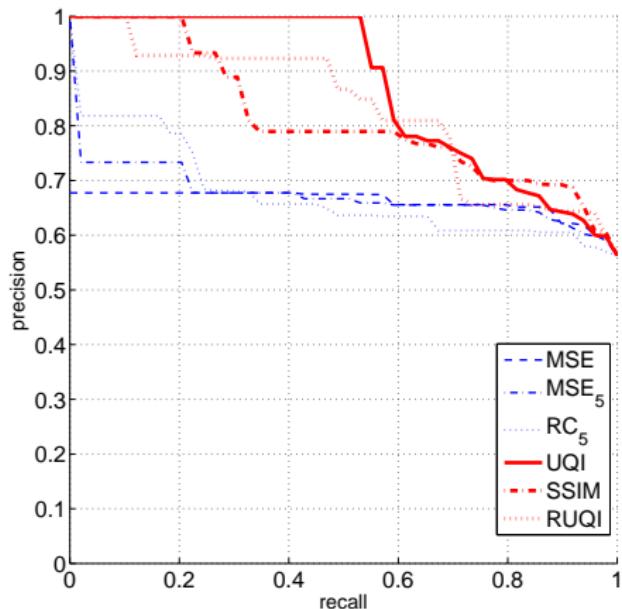


Résultats



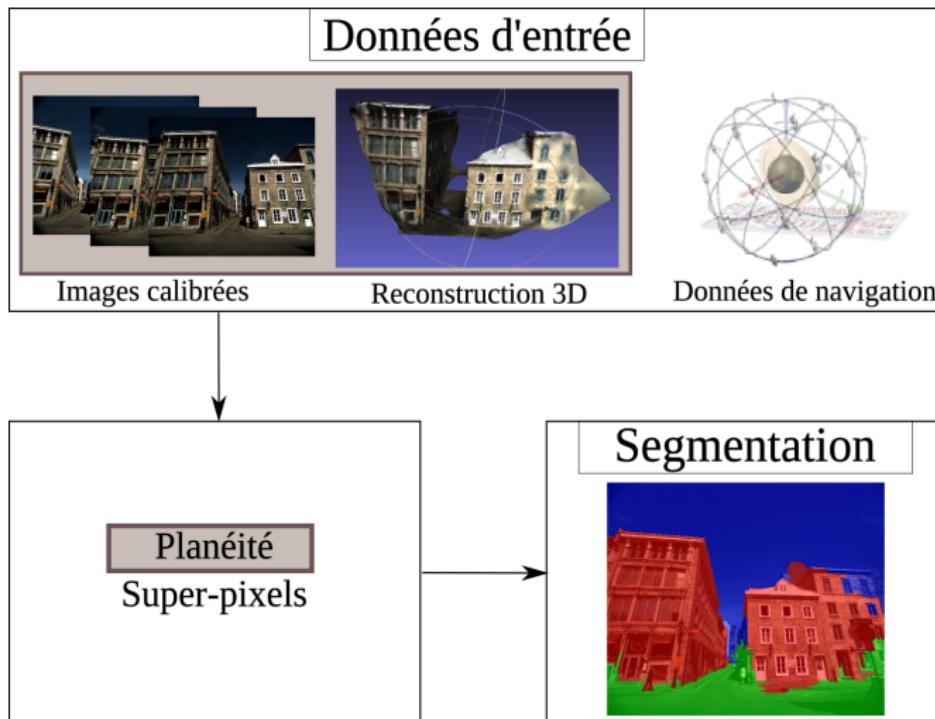
Courbe ROC (*Receiver Operator Characteristic*)

Résultats



Courbe PR (*Precision-Recall*)

Conclusion



Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

Super-pixels

Propriétés attendues

Variabilité des contraintes fixées lors de la construction :

- Propriétés d'apparence : contours, régions, couleurs, textures, etc.

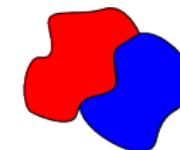
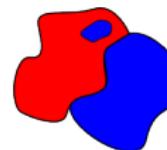
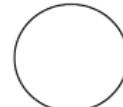
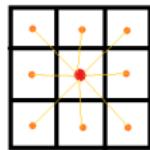


Super-pixels

Propriétés attendues

Variabilité des contraintes fixées lors de la construction :

- Propriétés d'apparence : contours, régions, couleurs, textures, etc.
- Propriétés spatiales : régularité topologique, compacité, connexité, etc.



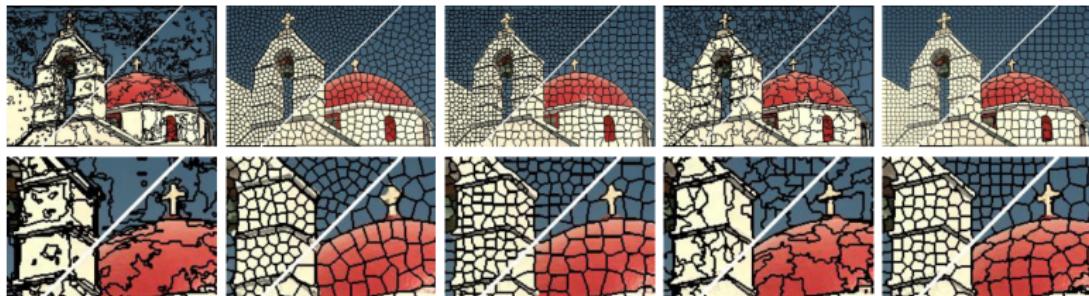
Super-pixels

Propriétés attendues

Variabilité des contraintes fixées lors de la construction :

- Propriétés d'apparence : contours, régions, couleurs, textures, etc.
- Propriétés spatiales : régularité topologique, compacité, connexité, etc.
- Propriétés d'invariance temporelle

Super-pixels



[Felzenszwalb, 2004, Shi, 2000, Levinshtein, 2009, Fulkerson, 2009, Achanta, 2012]

Algorithme SLIC [Achanta, 2012]

Algorithme SLIC - *Simple Linear Iterative Clustering*

Données : Image I , paramètres (K, m)

Résultat : Image segmentée $\{R_K\}$

$\{C_K\}^1 \leftarrow \text{InitCentres}(I, K)$

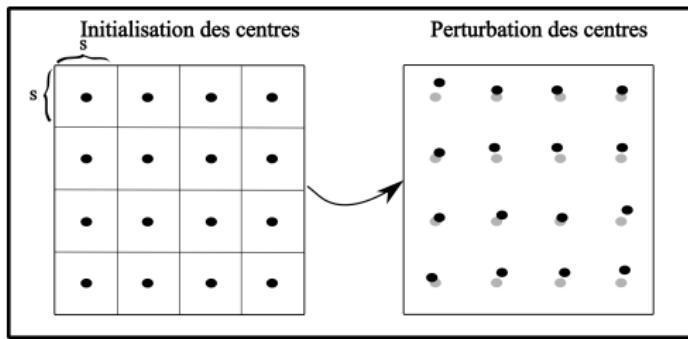
pour $i = 1 : N_{it}$ **faire**

$\{R_K\}^i \leftarrow \text{CalculRegions}(I, \{C_K\}^i, D_{SLIC})$
 $\{C_K\}^{i+1} \leftarrow \text{MiseAJourCentres}(\{R_K\}^i)$

fin

$\{R_K\} \leftarrow \text{RenforceConnexite}(\{R_K\}^{N_{it}})$

Algorithme SLIC



Algorithme SLIC - *Simple Linear Iterative Clustering*

Données : Image I , paramètres (K, m)

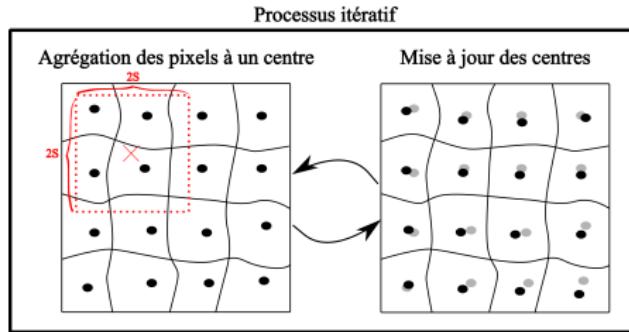
Résultat : Image segmentée $\{R_K\}$

```

 $\{C_K\}^1 \leftarrow \text{InitCentres}(I, K)$ 
pour  $i = 1 : N_{it}$  faire
   $\{R_K\}^i \leftarrow \text{CalculRegions}(I, \{C_K\}^i, D_{SLIC})$ 
   $\{C_K\}^{i+1} \leftarrow \text{MiseAJourCentres}(\{R_K\}^i)$ 
fin
 $\{R_K\} \leftarrow \text{RenforceConnexite}(\{R_K\}^{N_{it}})$ 

```

Algorithme SLIC



Algorithme SLIC - *Simple Linear Iterative Clustering*

Données : Image I , paramètres (K, m)

Résultat : Image segmentée $\{R_K\}$

$\{C_K\}^1 \leftarrow \text{InitCentres}(I, K)$

pour $i = 1 : N_{it}$ **faire**

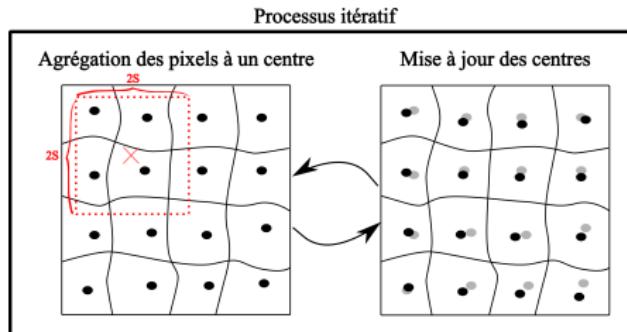
$\{R_K\}^i \leftarrow \text{CalculRegions}(I, \{C_K\}^i, D_{SLIC}, m)$

$\{C_K\}^{i+1} \leftarrow \text{MiseAJourCentres}(\{R_K\}^i)$

fin

$\{R_K\} \leftarrow \text{RenforceConnexite}(\{R_K\}^{N_{it}})$

Algorithme SLIC



Algorithme SLIC - *Simple Linear Iterative Clustering*

Données : Image I , paramètres (K, m)

Résultat : Image segmentée $\{R_K\}$

$\{C_K\}^1 \leftarrow \text{InitCentres}(I, K)$

pour $i = 1 : N_{it}$ **faire**

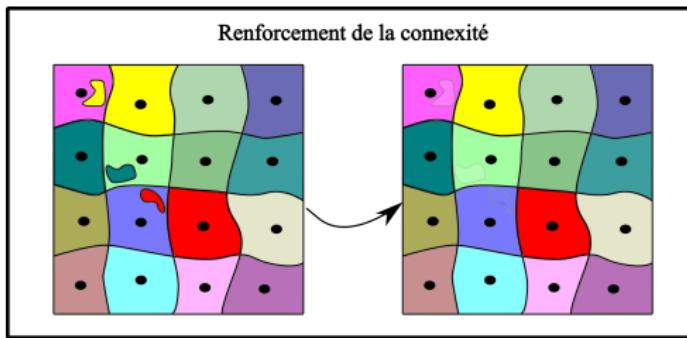
$\{R_K\}^i \leftarrow \text{CalculRegions}(I, \{C_K\}^i, D_{SLIC})$

$\{C_K\}^{i+1} \leftarrow \text{MiseAJourCentres}(\{R_K\}^i)$

fin

$\{R_K\} \leftarrow \text{RenforceConnexite}(\{R_K\}^{N_{it}})$

Algorithme SLIC



Algorithme SLIC - *Simple Linear Iterative Clustering*

Données : Image I , paramètres (K, m)

Résultat : Image segmentée $\{R_K\}$

$\{C_K\}^1 \leftarrow \text{InitCentres}(I, K)$

pour $i = 1 : N_{it}$ **faire**

$\{R_K\}^i \leftarrow \text{CalculRegions}(I, \{C_K\}^i, D_{SLIC})$
 $\{C_K\}^{i+1} \leftarrow \text{MiseAJourCentres}(\{R_K\}^i)$

fin

$\{R_K\} \leftarrow \text{RenforceConnexite}(\{R_K\}^{N_{it}})$

Algorithme SLIC

Distance d'agrégation

$$D_{SLIC}(p_j, p_i) = d_c(p_j, p_i) + \frac{m}{S} d_s(p_j, p_i) \quad (5)$$

p_i et p_j pixels étudiés, par rapport à un centre donné

- $d_c(p_j, p_i) = (I_j - I_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2$ distance couleur
- $d_s(p_j, p_i) = (x_j - x_i)^2 + (y_j - y_i)^2$ distance de position
- m , paramètre de compacité
- $S = \sqrt{N/K}$, constante spatiale

Algorithme SLIC

Extension de SLIC adaptatif : SLICO

$$D_{SLICO,n} = d_{c,n} + d_{s,n} \quad (6)$$

$$d_{c,n}(p_j, p_i) = \frac{d_{c,n}}{m_{c,n-1}} \text{ avec } d_{c,0} = d_{c_0}$$

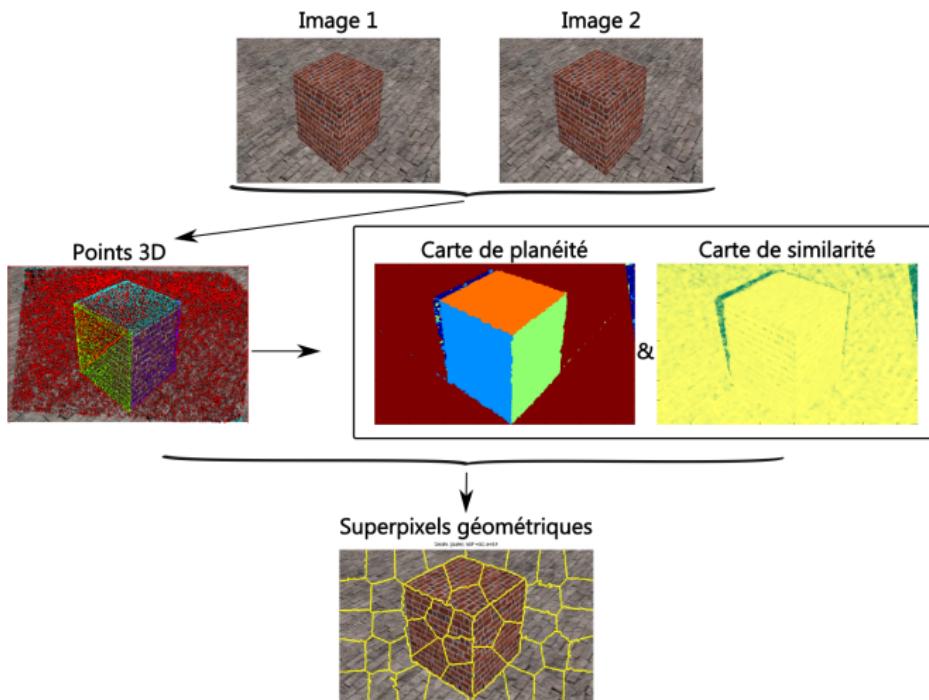
$$d_{s,n}(p_j, p_i) = \frac{d_{s,n}}{m_{s,n-1}} \text{ avec } d_{s,0} = d_{s_0} \quad (7)$$

$$m_{c,n-1} = \max(d_{c,n-1})$$

$$m_{s,n-1} = \max(d_{s,n-1})$$

où d_{c_0} et d_{s_0} sont les valeurs par défaut utilisées dans SLIC

Vue générale de l'approche de super-pixels proposée



Informations géométriques

Estimation des plans

Estimation des plans d'un nuage de points 3D avec une approche de type RANSAC [Fischer, 1981] :

- recouvrement de plans
- maximisation du nombre de points et minimisation de la distance moyenne des points au plan hypothétique
- détection multiple de plans, avec une approche itérative.

Informations géométriques

Estimation des plans

Estimation des plans d'un nuage de points 3D avec une approche de type RANSAC [Fischer, 1981] :

- recouvrement de plans
- maximisation du nombre de points et minimisation de la distance moyenne des points au plan hypothétique
- détection multiple de plans, avec une approche itérative.

Carte de planéité

→ étiquette d'appartenance à un plan 3D

$$d_{\vec{n}}(p_j, p_i) = \frac{1 + \cos(\vec{n}_i, \vec{n}_j)}{2}$$

Informations géométriques

Calcul de l'homographie

Estimation de l'image adjacente recalée par l'homographie induite par le plan de support

Informations géométriques

Calcul de l'homographie

Estimation de l'image adjacente recalée par l'homographie induite par le plan de support

Carte de similarité

→ quantification de la similarité inter-images

$$d_{UQI}(p_j, p_i) = UQI(p_j) \mathbb{1}_{UQI > \epsilon}$$

Distance d'agrégation proposée : D_{GEOM}

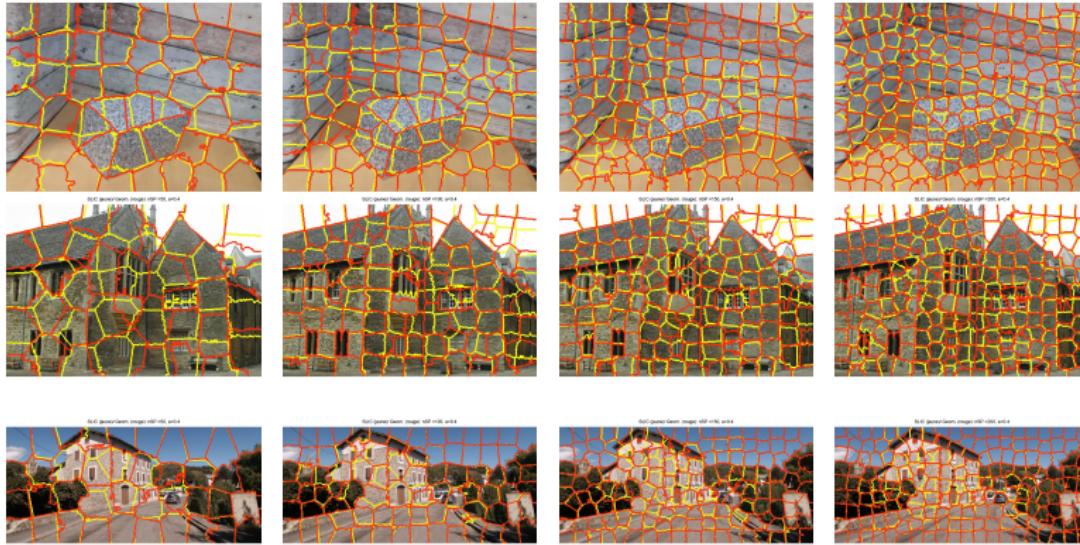
D_{GEOM}

$$D_{GEOM} = \sqrt{\lambda D_{SLICO} + (1 - \lambda) d_g} \quad (8)$$

Avec :

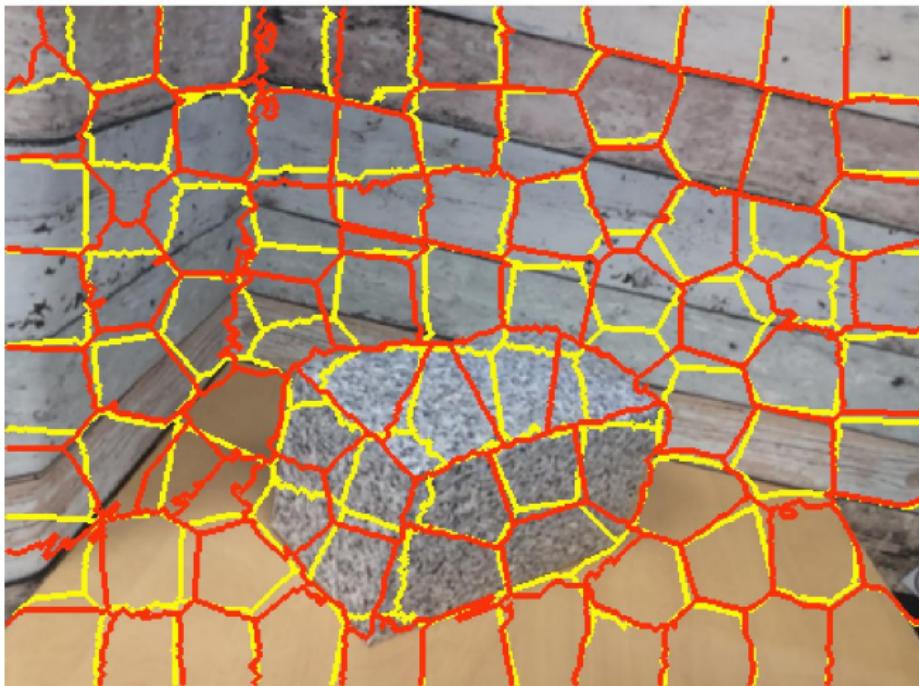
$$d_g(p_j, p_i) = 1 - d_{\vec{n}}(p_j, p_i) \ d_{UQI}(p_j) \quad (9)$$

Résultats



SLICO (en jaune) et GEOM (en rouge), pour 50, 100, 150 et 200 super-pixels

Résultats



Résultats

SLIC (jaune)/ Geom. (rouge): nSP =100, a=0.4



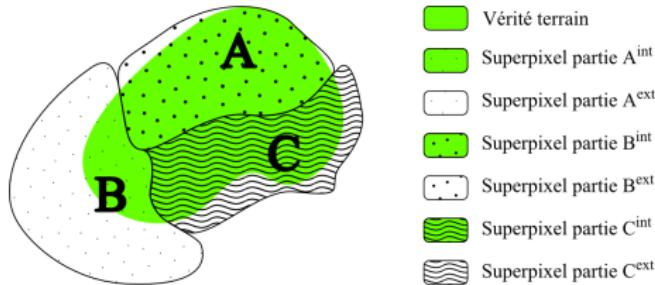
Résultats

SLIC (jaune)/ Geom. (rouge): nSP =100, a=0.4



Évaluation d'une sur-segmentation

Comparaison du résultat à une vérité terrain, ou segmentation de référence par une métrique



Erreur de sous-segmentation

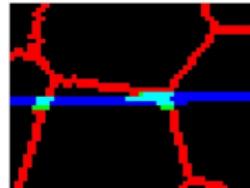
$$ErrSS = \frac{1}{N} \left[\sum_{i=1}^M \left(\sum_{s_j | s_j \cap g_i > B} \#s_j \right) - N \right] \quad (10)$$

Évaluation locale

Distance aux contours

$$m_{\text{dContours}} = \frac{1}{N_{\text{VT}}} \sum_{p_i \in \text{VT}} d_i(p_i) \quad (11)$$

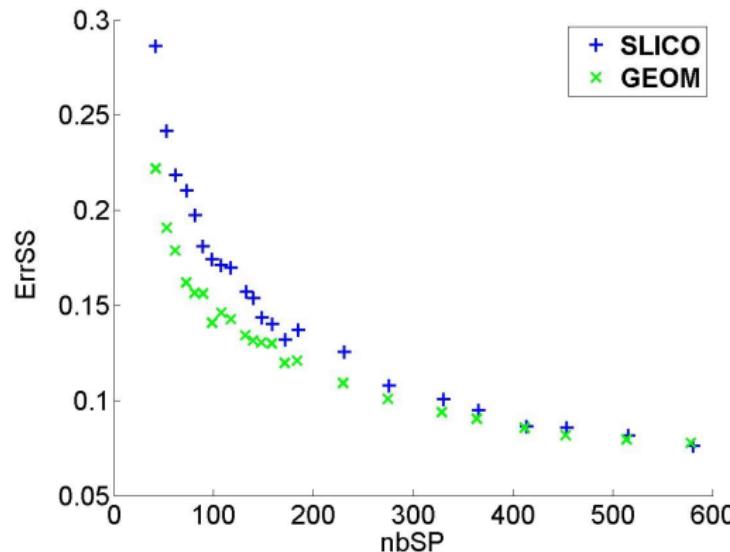
- $d_i(p_i)$, distance du pixel p_i appartenant à la vérité terrain (VT), au contour le plus proche de la sur-segmentation
- N_{VT} , nombre de pixels sur le contour de la vérité terrain



Analyse globale

Sensibilité aux paramètres

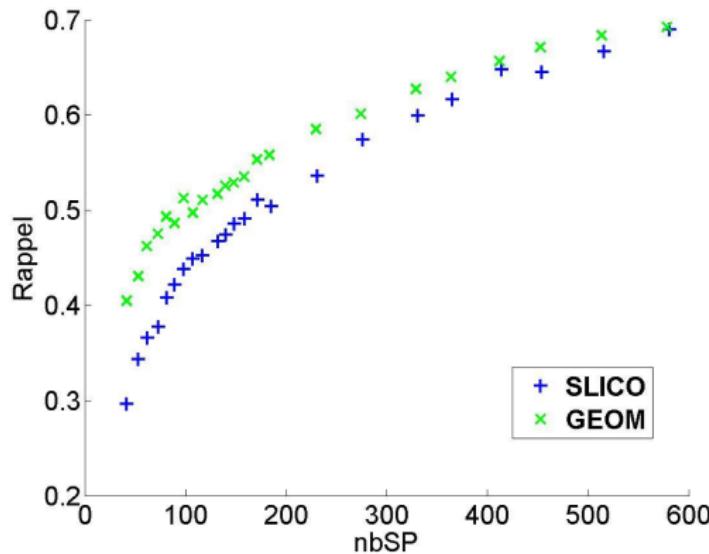
Trois paramètres : nombre SP , ϵ , λ



Analyse globale

Sensibilité aux paramètres

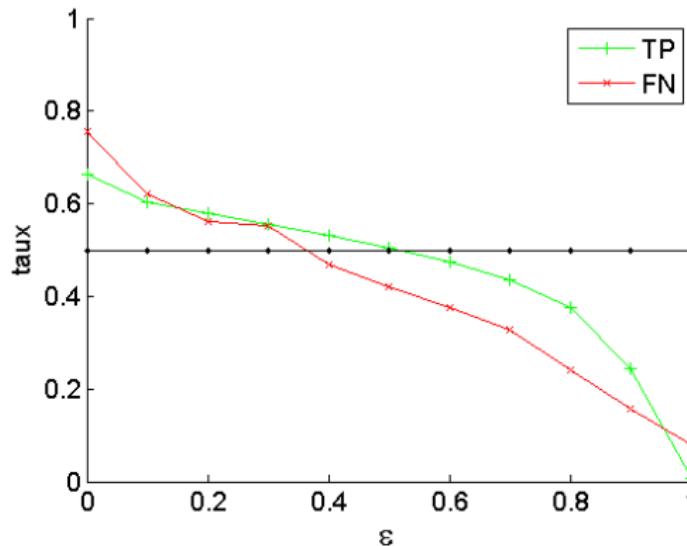
Trois paramètres : nombre SP, ϵ , λ



Analyse globale

Sensibilité aux paramètres

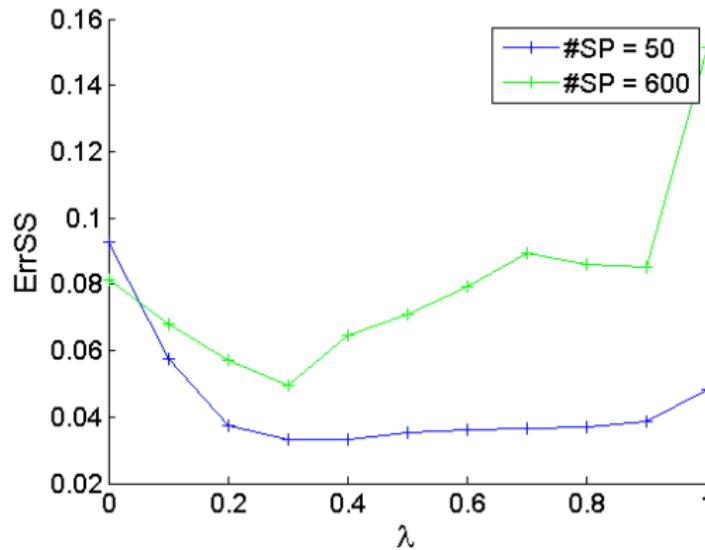
Trois paramètres : nombre SP, ϵ , λ



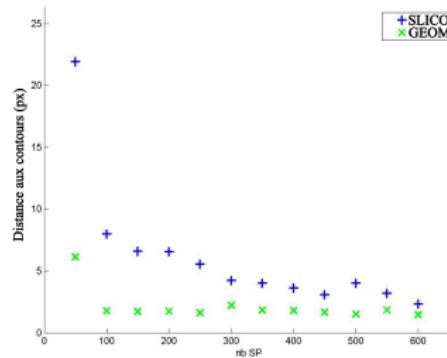
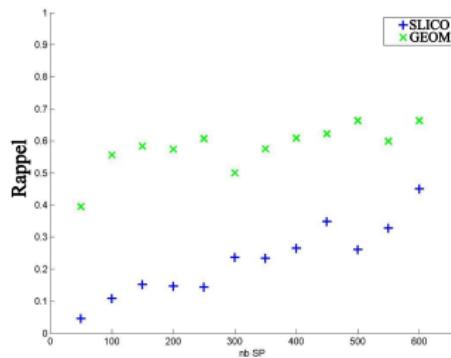
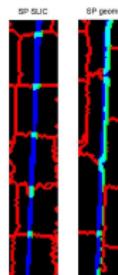
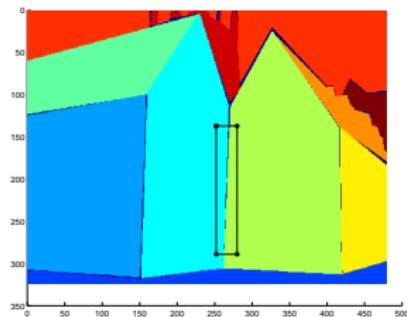
Analyse globale

Sensibilité aux paramètres

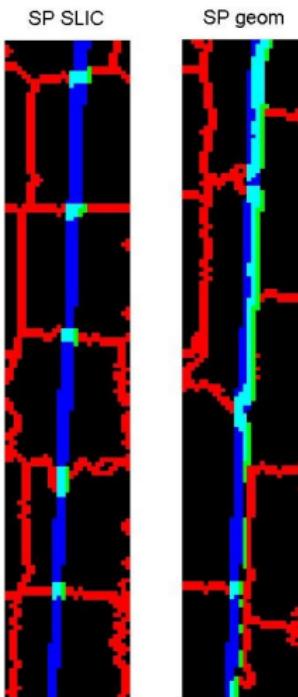
Trois paramètres : nombre SP, ϵ , λ



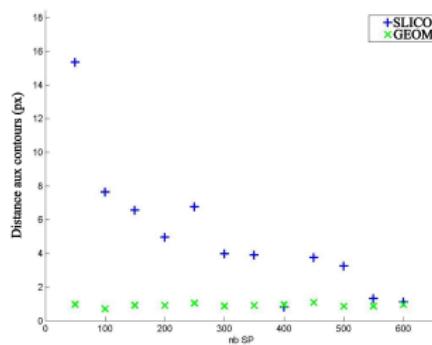
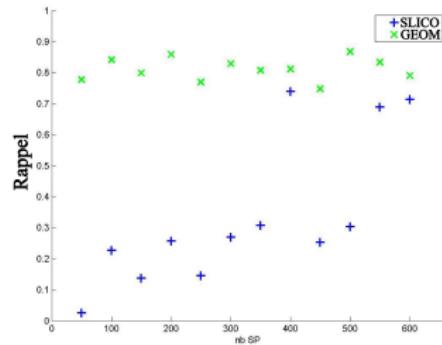
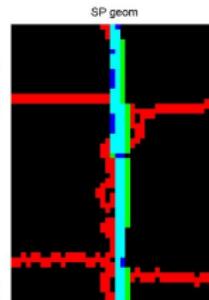
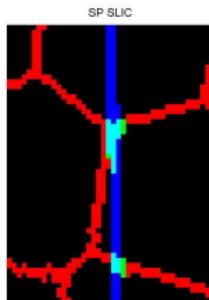
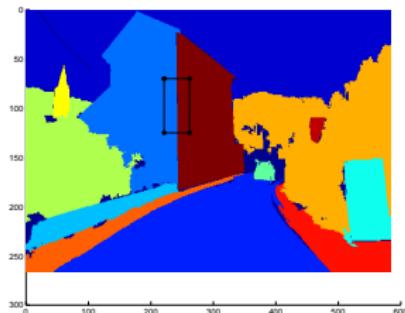
Analyse locale : Exemple 1



Analyse locale : Exemple 1



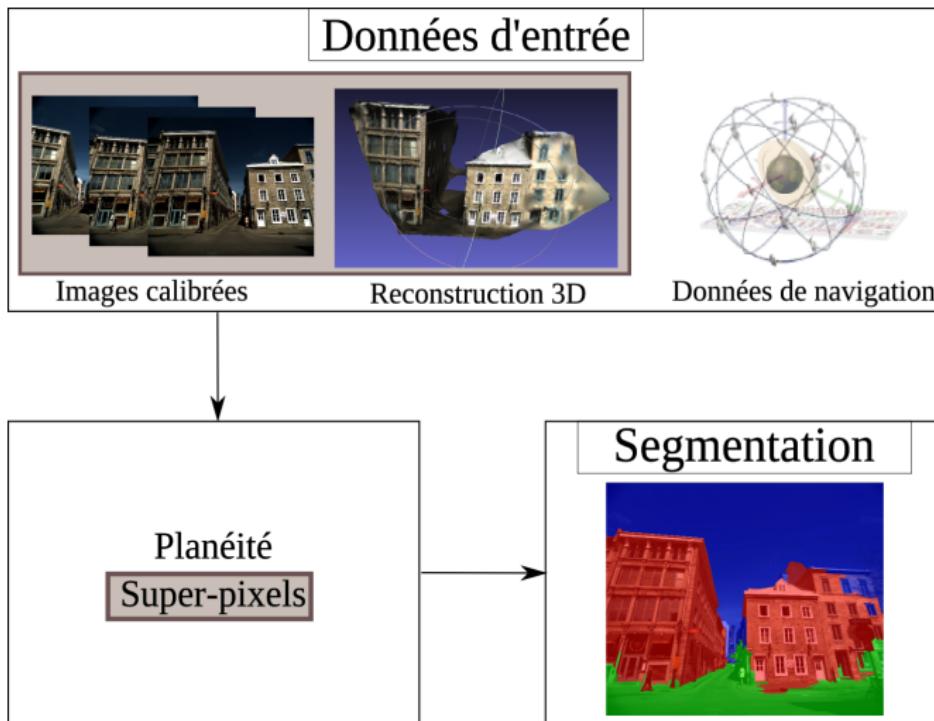
Analyse locale : Exemple 2



Conclusion

- Constructeur de super-pixels
- Combinaison de l'information géométrique et photométrique pour une sur-segmentation plus robuste
- Performance supérieure à SLIC/SLICO (erreur de sous-segmentation, rappel global et local, distance aux contours)

Conclusion



Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

Vue générale de l'approche proposée

Données d'entrée (images, paramètres de calibrage et nuage de points 3D maillé)



Vue générale de l'approche proposée

Données d'entrée (images, paramètres de calibrage et nuage de points 3D maillé)



Estimation des plans 3D

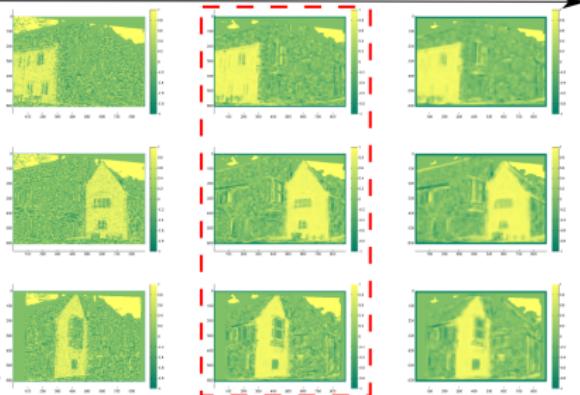


Vue générale de l'approche proposée

Calcul de la similarité inter-images pour la classification des pixels en zones planaires

Echelles

Plans



Résultat de la classification en zones planaires

Vue générale de l'approche proposée

Construction des superpixels géométriques

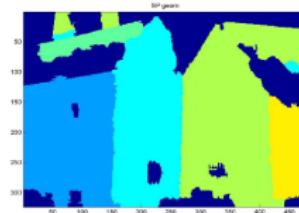


Vue générale de l'approche proposée

Construction des superpixels géométriques



Segmentation sémantique



Analyses

Sensibilité aux types de reconstruction

- Nuage épars (E) SfM (VisualSfM [Wu, 2011])
- Nuage dense (D) MVS (CM3MVS [Jancosek, 2011])

Analyses

Sensibilité aux types de reconstruction

- Nuage épars (E) SfM (VisualSfM [Wu, 2011])
- Nuage dense (D) MVS (CMPMVS [Jancosek, 2011])

	MEE	MDD	MDE
Estimation des plans (points 3D)	épars	dense	dense
Calcul des homographies (points 2D)	épars	dense	épars

Analyses

Sensibilité aux types de reconstruction

- Nuage épars (E) SfM (VisualSfM [Wu, 2011])
- Nuage dense (D) MVS (CM3PMVS [Jancosek, 2011])

	MEE	MDD	MDE
Estimation des plans (points 3D)	épars	dense	dense
Calcul des homographies (points 2D)	épars	dense	épars

Analyses

Sensibilité aux types de reconstruction

- Nuage épars (E) SfM (VisualSfM [Wu, 2011])
- Nuage dense (D) MVS (CM3PMVS [Jancosek, 2011])

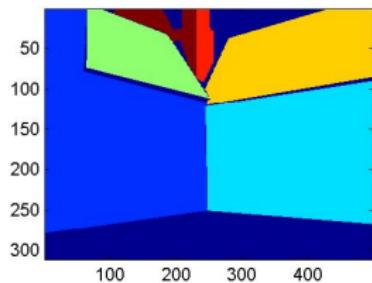
	MEE	MDD	MDE
Estimation des plans (points 3D)	épars	dense	dense
Calcul des homographies (points 2D)	épars	dense	épars

Influence de l'échelle sur la similarité inter-images

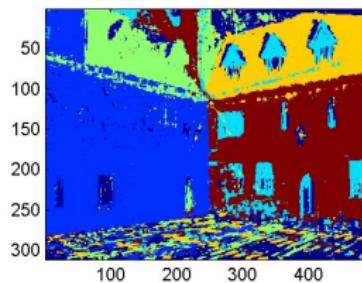
→ *animation3_plane2_UQI.gif*

Résultats : Exemple 1

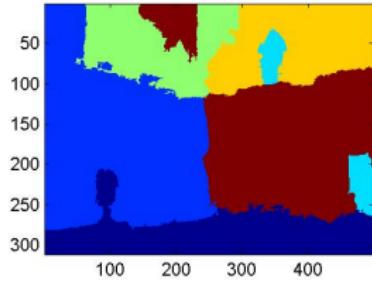
VT



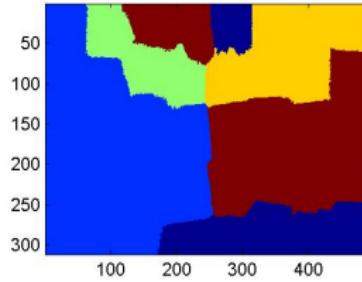
Pixel



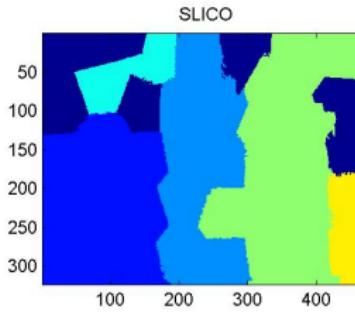
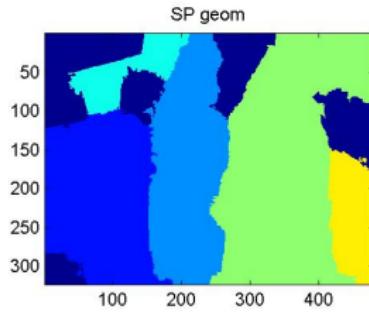
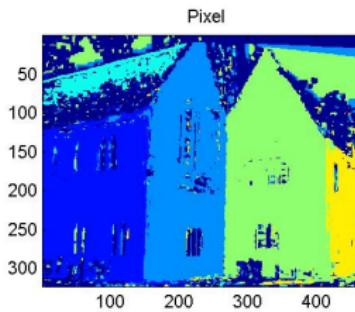
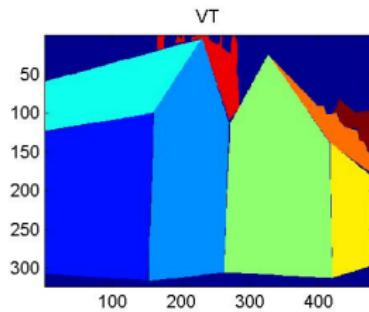
SP geom



SLICO



Résultats : Exemple 2



Résultats

Approche	taux de TP	taux de F
PIXEL	81.21 [57.4,97.9]	7.40 [0.5,16.2]
SLICO	83.70 [57.3,96.9]	5.79 [2.2,11.0]
GEOM	87.28 [59.8,98.1]	4.73 [0.4,9.7]

Résultats

Résultats visuels avec [Bódis-Szomorú, 2014]



Résultats

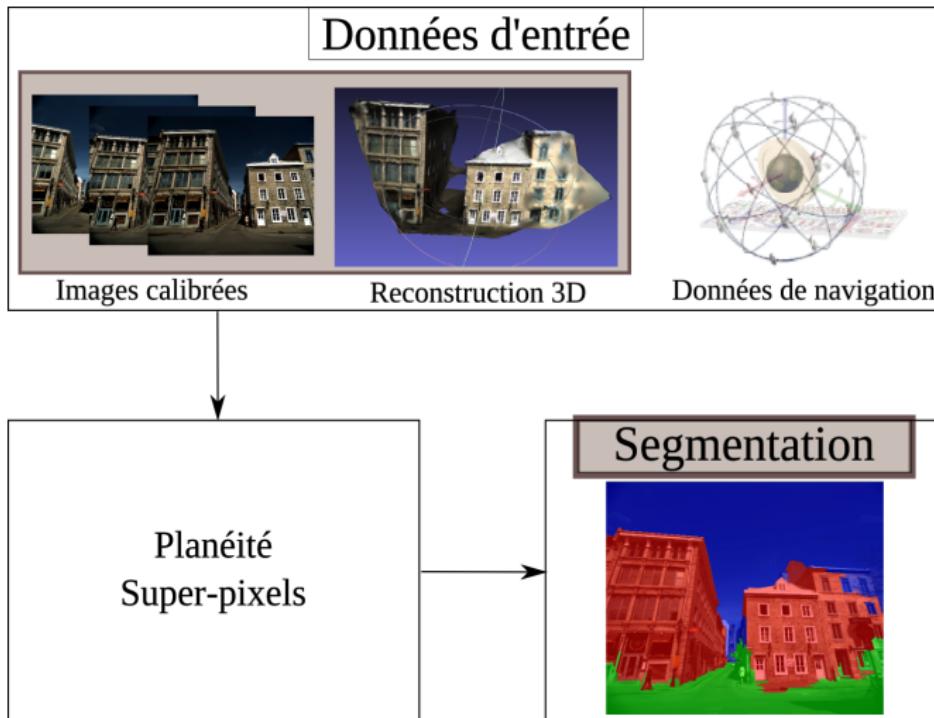
Résultats visuels avec [Bódis-Szomorú, 2014]



Résultats visuels avec notre approche



Conclusion



Conclusion

- Proposition d'une approche de segmentation sémantique en plans
- Fusion de l'information 2D (similarité inter-images)/3D (information géométrique)
- Utilisation des super-pixels intégrant de l'information géométrique (normale, planéité)

Sommaire

- 1 Travaux existants
- 2 Méthodologie
- 3 Mesures de cohérence photométrique inter-images et planéité
- 4 Super-pixels et géométrie
- 5 Segmentation sémantique en plans
- 6 Conclusion et perspectives

Conclusion

- Utilité de la similarité inter-images
- Protocole d'évaluation pour les mesures IQA, pour la classification des ZP/ZNP
- Proposition d'un constructeur de super-pixels intégrant la géométrie
- Proposition d'une approche de segmentation sémantique en plans

Conclusion

- Utilité de la similarité inter-images → ORASIS
- Protocole d'évaluation pour les mesures IQA, pour la classification des ZP/ZNP → ICPRAM
- Proposition d'un constructeur de super-pixels intégrant la géométrie → VISAPP
- Proposition d'une approche de segmentation sémantique en plans

→ 2 articles en conférences internationales + 1 nationale

Perspectives

- Utilisation de l'information sur l'orientation des lignes
- Segmentation 3D : Super-pixels 3D \leftrightarrow super-voxels
- Intégration d'informations *a priori* : supervisée ou utilisateurs

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.
- [Alvarez, 2008] Alvarez, J. and Gevers, T. and LeCun, Y. and López, A., Road Scene Segmentation from a Single Image, *IEEE ECCV*, 2008.
- [Bartoli, 2006] A. Bartoli, A Random Sampling Strategy for piecewise Planar Scene Segmentation, *CVIU*, 2006.
- [Bódis-Szomorú, 2014] A. Bódis-Szomorú, H. Riemenschneider and L. Van Gool, Fast, Approximate Piecewise-Planar Modeling Based on Sparse Structure-from-Motion and Superpixels, *CVPR*, 2014.
- [Badrinarayanan, 2015] Badrinarayanan, V. and Handa, A. and Cipolla, R., SegNet : A Deep Convolutional Encoder-Decoder Architecture for Robust Semantic Pixel-Wise Labelling, *arXiv preprint*, 2015.
- [Brostow, 2008] G. Brostow and J. Shotton and J. Fauqueur and R. Cipolla, Segmentation and Recognition Using Structure from Motion Point Clouds, *IEEE ECCV*, 2008.
- [Comaniciu, 2002] D. Comaniciu and P. Meer, Mean shift : a robust approach toward feature space analysis, *IEEE Transactions on PAMI*, pp. 603-619, 2002.
- [Dai, 2013] J. Dai, Y.N. Wu, J. Zhou and S. Zhu, Cosegmentation and Cosketch by Unsupervised Learning, *IEEE ICCV*, 2013.
- [Delage, 2007] Delage, E. and Lee, H. and Ng, A., Automatic Single-Image 3D Reconstructions of Indoor Manhattan World Scenes, *Robotics Research*, 2007.
- [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.

References II

- [Fouhey, 2013] Fouhey, D. and Gupta, A. and Hebert, M., Data-Driven 3D Primitives for Single Image Understanding, *IEEE ICCV*, 2013.
- [Fulkerson, 2009] B. Fulkerson and A. Vedaldi and S. Soatto, Class segmentation and object localization with superpixel neighborhoods, *ICCV*, 2009.
- [Gallup, 2008] D. Gallup and J.-M. Frahm and M. Pollefeys, Piecewise planar and non-planar stereo for urban scene reconstruction, *IEEE CVPR*, 2010.
- [Grundmann, 2010] 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.
- [Hoiem, 2011] Hoiem, D. and Efros, A.A. and Hebert, M., Recovering occlusion boundaries from an image, *IEEE IJCV*, 2011.
- [Jancosek, 2011] Jancosek, M. and Pajdla, T., Multi-view Reconstruction Preserving Weakly-supported Surfaces, *IEEE CVPR*, 2011.
- [Joulin, 2012] A. Joulin and F. Bach and J. Ponce, Multi-Class Cosegmentation *IEEE CVPR*, 2012.
- [Kutulakos, 2000] K. Kutulakos, Approximate N-View Stereo, *IEEE ECCV*, 2000.
- [Levinshtein, 2009] A. Levinshtein and A. Stere and K. N. Kutulakos and D. J. Fleet and S. J. Dickinson and K. Siddiqi, TurboPixels : Fast Superpixels Using Geometric Flows *IEEE Transactions on PAMI*, 2009.
- [Raza, 2013] S. H. Raza and M. Grundmann and I. Essa, Geometric Context from Video, *IEEE CVPR*, 2013.
- [Shi, 2000] J. Shi and J. Malik, Normalized cuts and image segmentation, *IEEE Transactions on PAMI*, 2000.
- [Sinha, 2008] S.N. Sinha, D. Steedly and R. Szeliski, Interactive 3D architectural modeling from unordered photo collections, *ACM Transactions on Graphic*, 2008.

References III

- [Sinha, 2014] S.N. Sinha and D. Scharstein and R. Szeliski, Efficient High-Resolution Stereo Matching using Local Plane Sweeps, *IEEE CVPR*, 2014.
- [Tighe, 2013] Tighe, J. and Lazebnik, S., Superparsing - Scalable Nonparametric Image Parsing with Superpixels, *IJCV*, 2013.
- [Wang, 2002] Z. Wang and A.C. Bovik, A universal image quality index, *IEEE Signal Processing Letters*, 2002.
- [Wang, 2004] Z. Wang and A.C. Bovik and H.R. Sheikh and E.P. Simoncelli, Image Quality Assessment : From Error Visibility to Structural Similarity, *IEEE Transaction on Image Processing*, 2004.
- [Wu, 2011] C. Wu, VisualSFM : A Visual Structure from Motion System, 2011.

Résultats Segmentation en 3 classes

