

Traitement joint de nuages de points et de photos pour l'acquisition de scènes urbaines

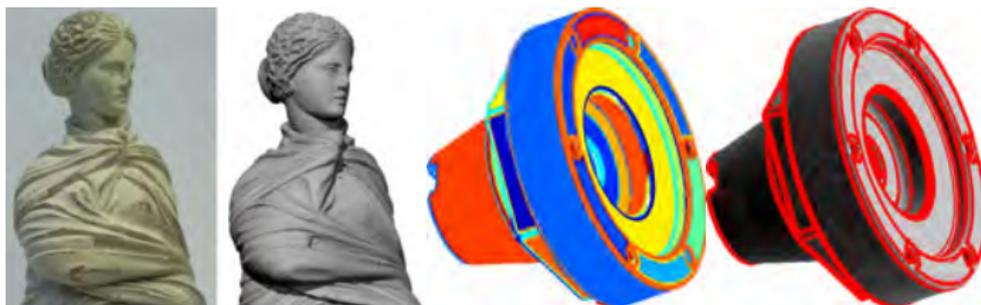
Maximilien Guislain Julie Digne Raphaëlle Chaine

LIRIS - Équipe GeoMod - CNRS en collaboration avec Technodigit



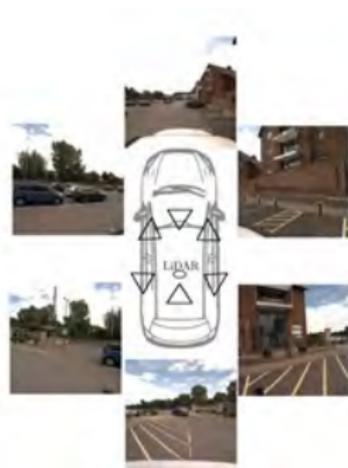
Un peu d'archéologie...

- Le projet "Géométrie Inverse pour l'Industrie" (LURPA - CMLA)



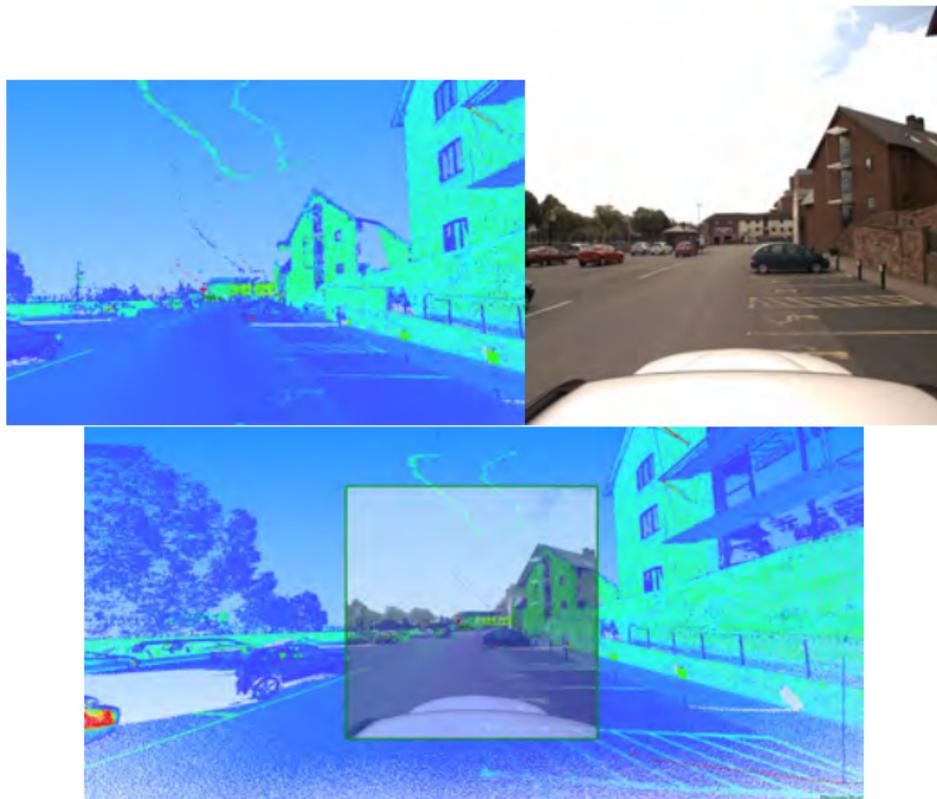
Acquisition de données urbaines

- Acquisition d'une scène urbaine sous forme d'un *nuage de points* et d'un ensemble de *photos*.
- *Nuage de points*: un scanner laser monté sur véhicule en mouvement.
- *Photos*: à intervalles réguliers des photos sont prises par des appareils photos montés sur le véhicule.



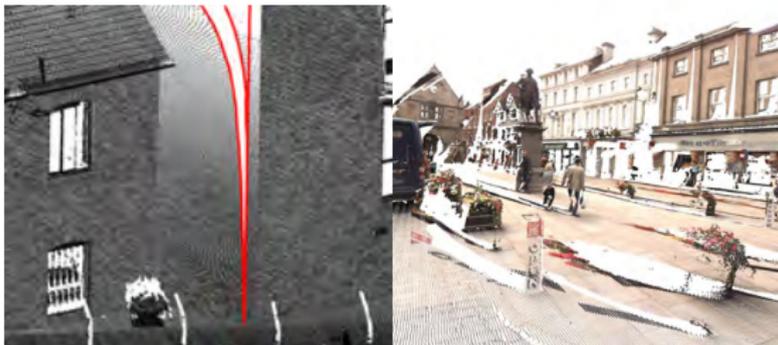
Projet mené en collaboration avec l'entreprise Technodigit (thèse CIFRE de Maximilien Guislain)

Quelques défis de ces données



Masse de données

Quelques défis de ces données



Transparence, mauvais recalage

Un outil pour le recalage: Information Mutuelle

Principe

Vérifier que deux images (ou plus généralement deux signaux) encodent bien la même information.

- $p(X = x) = \frac{\text{Nbre de pixels ayant la valeur } x}{\text{Nbre de pixels de l'image}}$
- Probabilité jointe:

$$P(x, y) = P(\text{Un pixel prend la valeur } x \text{ dans } I \text{ et } y \text{ dans } J)$$

Un outil pour le recalage: Information Mutuelle

Principe

Vérifier que deux images (ou plus généralement deux signaux) encodent bien la même information.

- Utile dans le cas de l'imagerie multi-modale
- $p(X = x) = \frac{\text{Nbre de pixels ayant la valeur } x}{\text{Nbre de pixels de l'image}}$
- Probabilité jointe:

$$P(x, y) = P(\text{Un pixel prend la valeur } x \text{ dans } I \text{ et } y \text{ dans } J)$$

Définitions

Information mutuelle

$$MI(X, Y) = \sum_x \sum_y p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right)$$

Définitions

Information mutuelle

$$MI(X, Y) = \sum_x \sum_y p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right)$$

- Si $I = J$ ou $I = f(J)$, $p(x, y) = p(x) = p(y)$ donc:

$$MI(X, Y) = -\sum_x p(x) \log p(x)$$

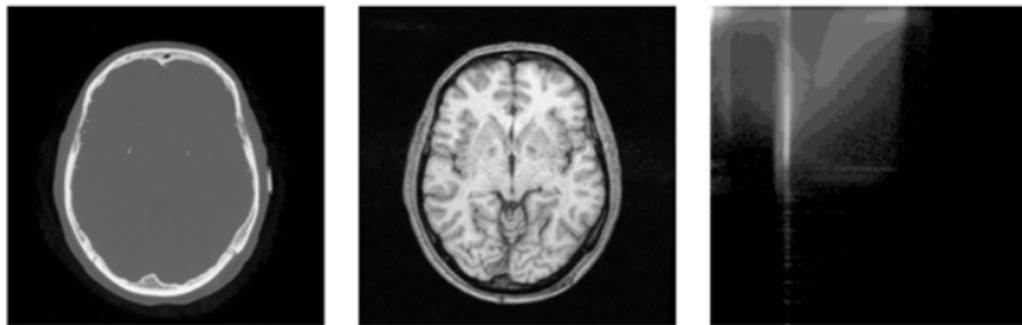
Définitions

Information mutuelle

$$MI(X, Y) = \sum_x \sum_y p(x, y) \log\left(\frac{p(x, y)}{p(x)p(y)}\right)$$

- Si $I = J$ ou $I = f(J)$, $p(x, y) = p(x) = p(y)$ donc:
 $MI(X, Y) = -\sum_x p(x) \log p(x)$
- I et J indépendantes: $p(x, y) = p(x)p(y)$ donc: $MI(X, Y) = 0$

Imagerie médicale



Pluim et al. (survey) 2003

- Intérêt: adaptée à l'imagerie multi-modale

Application de l'information mutuelle: le recalage de photos sur des formes

- Acquisition d'un objet 3D et prise de photos du même objet

Application de l'information mutuelle: le recalage de photos sur des formes

- Acquisition d'un objet 3D et prise de photos du même objet
- *i.e.* Comment savoir de quelle position a été prise la photo (*pose estimation*)

Application de l'information mutuelle: le recalage de photos sur des formes

- Acquisition d'un objet 3D et prise de photos du même objet
- *i.e.* Comment savoir de quelle position a été prise la photo (*pose estimation*)

Application de l'information mutuelle: le recalage de photos sur des formes

- Acquisition d'un objet 3D et prise de photos du même objet
- *i.e.* Comment savoir de quelle position a été prise la photo (*pose estimation*)

Principe

On considère le problème du recalage image-forme en une suite d'itérations du recalage image-image

Principe du recalage d'images sur des formes 3D

- À partir de la pose \mathcal{T} et de la forme S on peut rendre une image synthétique.

Principe du recalage d'images sur des formes 3D

- À partir de la pose \mathcal{T} et de la forme S on peut rendre une image synthétique.
- Cette image correspond à ce qu'une caméra verrait depuis la pose \mathcal{T}

Principe du recalage d'images sur des formes 3D

- À partir de la pose \mathcal{T} et de la forme S on peut rendre une image synthétique.
- Cette image correspond à ce qu'une caméra verrait depuis la pose \mathcal{T}
- Si la pose correspond à la pose qui a générée la photo I , I et I_S doivent représenter la même chose.

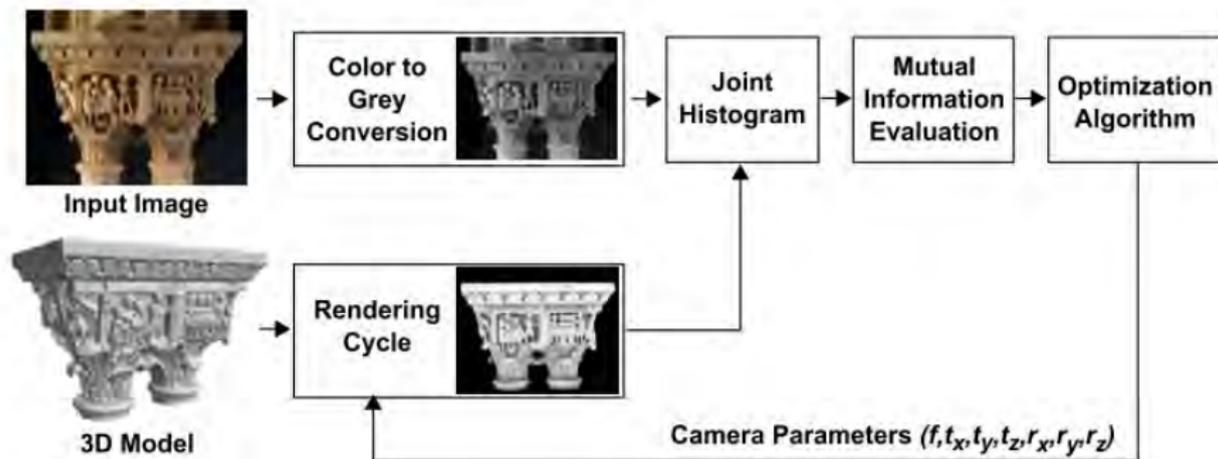
Principe du recalage d'images sur des formes 3D

- À partir de la pose \mathcal{T} et de la forme S on peut rendre une image synthétique.
- Cette image correspond à ce qu'une caméra verrait depuis la pose \mathcal{T}
- Si la pose correspond à la pose qui a générée la photo I , I et I_S doivent représenter la même chose.

Lien avec MI

L'information mutuelle entre I et I_S doit donc être maximale

Application de l'information mutuelle: le recalage de photos sur des formes



Dellepiane et al.

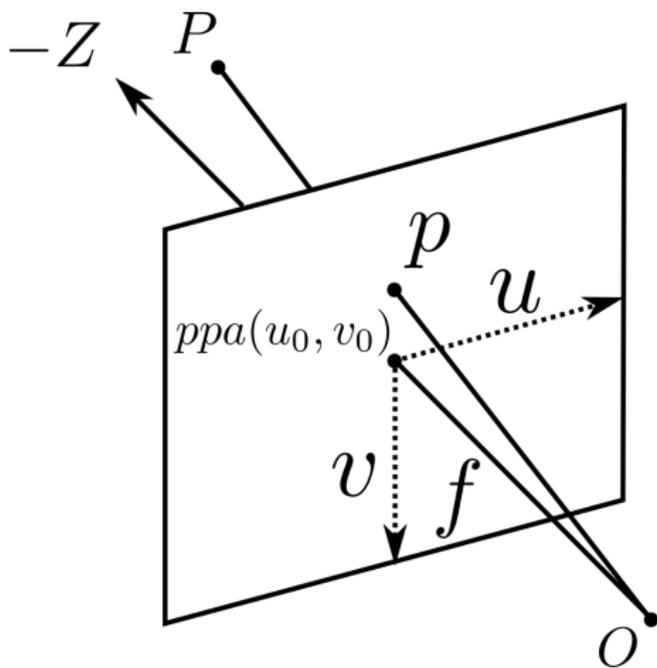
Résultats



Sottile et al. 2010 Mutual Correspondences

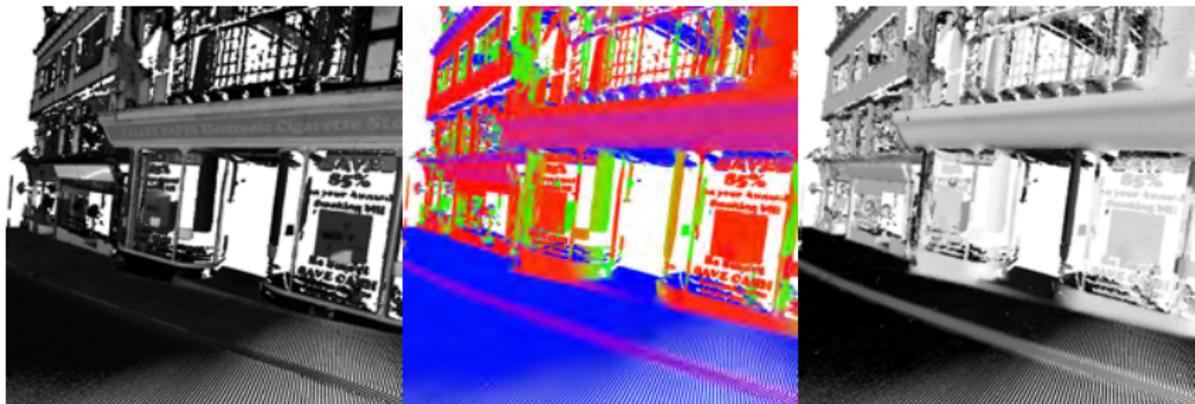
- La convergence est très lente
- Minima locaux \Rightarrow bonne pose initiale \mathcal{T}_0
- Mesure globale sur l'image et non locale.

Génération d'images synthétiques

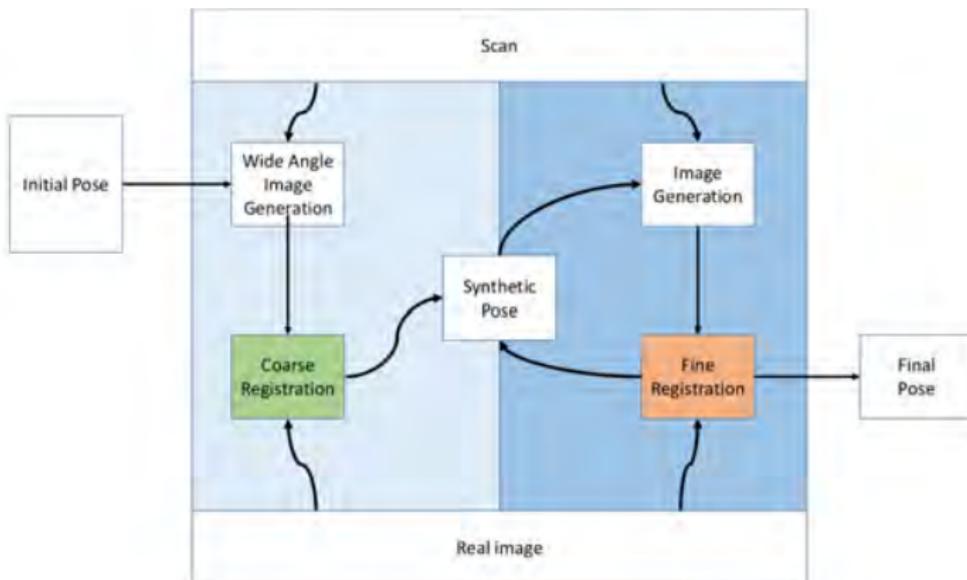


- Pour une pose donnée de l'appareil photo, on peut calculer un *rendu* du nuage de points.
- Quelle couleur donner à chaque point? Normales, intensité de retour laser...

Génération d'images synthétiques

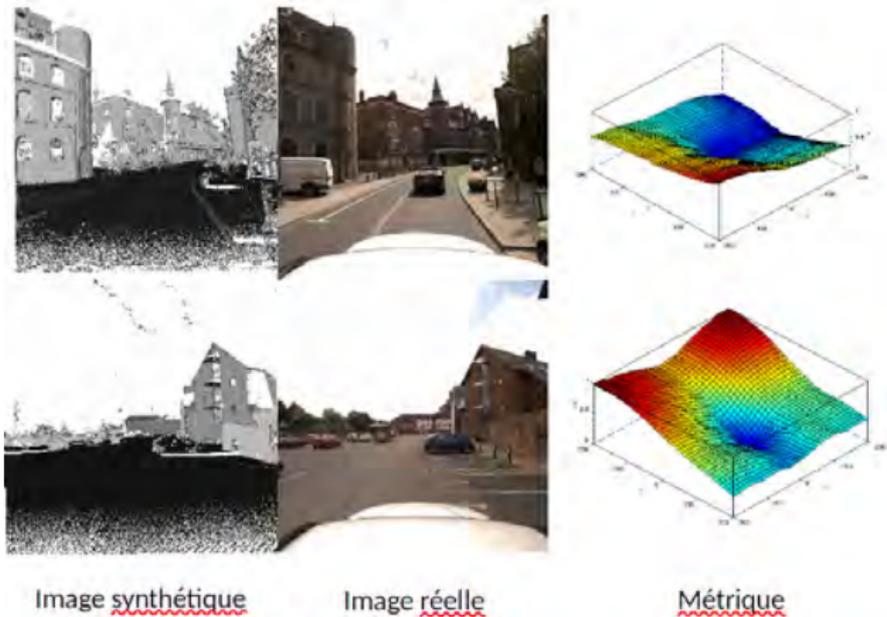


Méthode proposée



- Modification de l'information mutuelle pour prendre en compte la spatialisation

Adaptation de l'information mutuelle



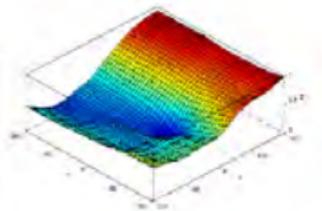
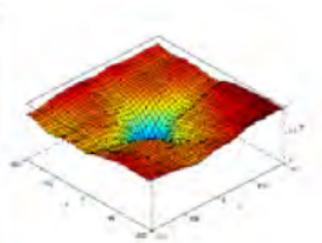
Normalized Mutual Information

Adaptation de l'information mutuelle



Image synthétique

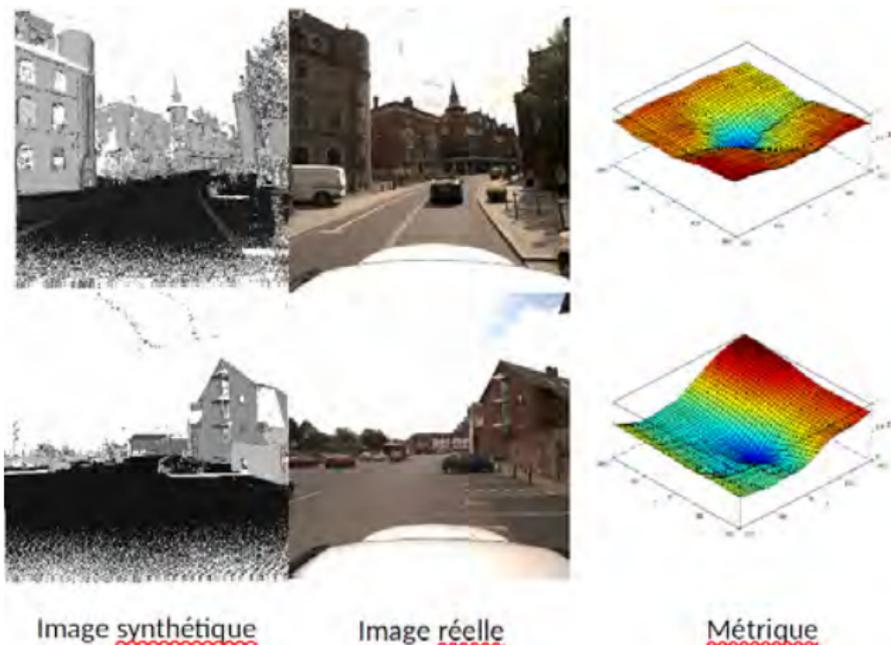
Image réelle



Métrique

Difference of Histograms of Oriented Gradients (DHOG)

Adaptation de l'information mutuelle



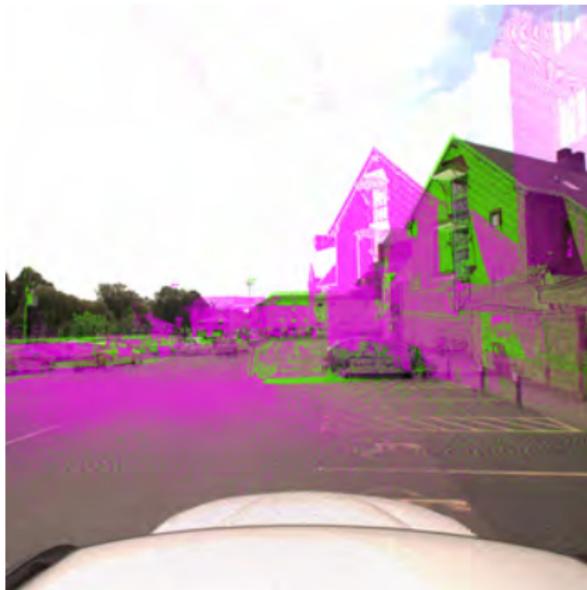
MIDHOG: Combination of both $MIDHOG = (2 - NMI) + \alpha DHOG$

Recalage grossier

- But: robustesse et rapidité
- Seule la rotation est estimée.
- Rotations \approx translations dans le plan image



Recalage grossier



Perturbation en azimuth -10°



Perturbation en azimuth $+15^\circ$

Recalage fin

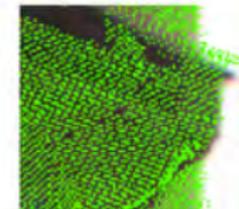
- But: précision
- Estimation de la rotation et translation



Résultats



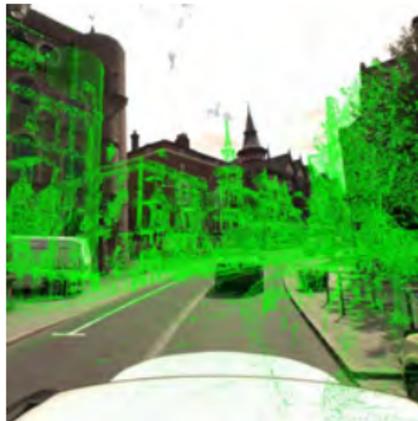
Recalage fin



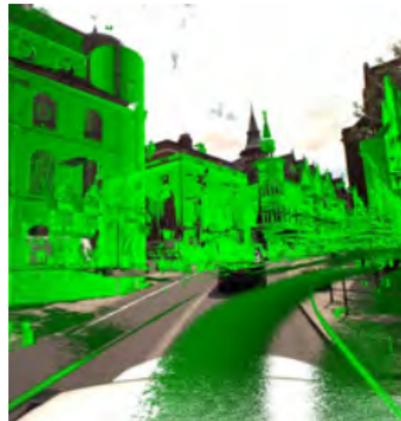
Résultats - comparaison



Ours



NMI

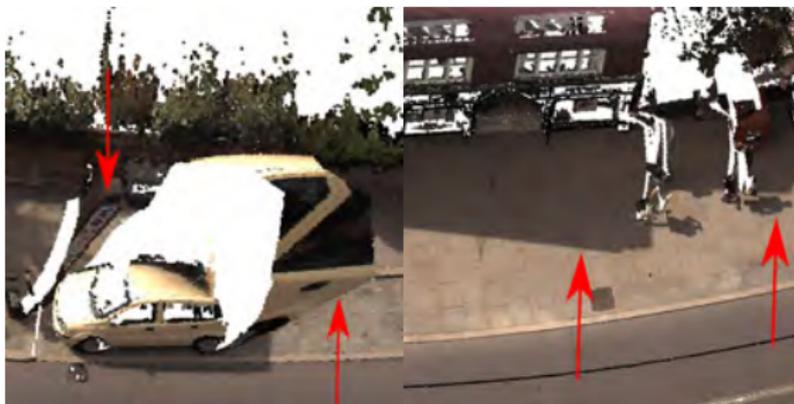


GOM

- Recalage fait sur un rendu basé normales.

Exploitation du recalage: colorisation

- Donner une couleur à chaque point du nuage de points



Les méthodes naïves échouent



- Problèmes similaires pour la colorisation basée sur la photo la plus parallèle aux murs, la plus proche, moyenne des couleurs...

Solution: couleur consensus par voisinage [Cho et al. 2014]

- Création d'un graphe couvrant le nuage de points (Delaunay, Z-ordering, nearest neighbor graph)
- Minimisation d'une énergie sur ce graphe:

$$E(C) = \sum_{p \in \mathcal{P}} D_p(C) + \lambda \sum_{p, q \in \mathcal{N}} S_{p, q}(C)$$

- $D_p(C)$ mesure la cohérence de la colorisation C avec le point p :
 $D_p(C) = \|C(p) - \text{med}(C(p))\|$
- $S_{p, q}$ mesure la cohérence avec la couleur des voisins
 $S_{p, q} = \|C(p) - C(q) - \text{med}(C(p) - C(q))\|$

Résultats



Résultats



Résultats

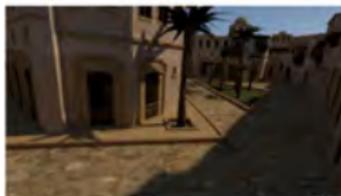


Résultats



Retrait des ombres portées

- Les ombres portées figent le rendu à une heure précise, c'est une information extrinsèque à la scène.
- Domaine de recherche des "Images intrinsèques": ici on va (beaucoup) simplifier le modèle.



Input I



Illumination L



Reflectance ρ

Images de Bonneel et al. 2017

Principe de la méthode



(a) Detected points.

(b) Filtered points.



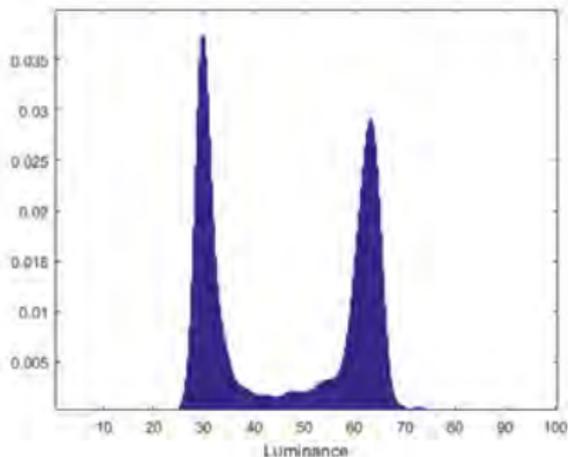
(c) Shadow interface points.

(d) Shadow mask.

Caractérisation des frontières ombres/lumières

Observation

À l'interface entre une zone d'ombre et une zone éclairée, la réflectance (intensité de retour) du laser ne varie pas alors que la luminance (L dans l'espace Lab) présente un histogramme bimodal.



Extrapolation sur l'ensemble de l'image

Formulation

L'extrapolation se fait dans le domaine image par minimisation d'énergie.

$$E(\delta) = \sum_{p: \text{image pixel}} E_{data}(\delta, p) + \gamma \sum_{(p, q): \text{pixels voisins}} E_{smooth}(\delta, p, q) \quad (1)$$

avec $E_{smooth} = e^{-\frac{1}{2\sigma^2}(L(p)-L(q))} \mathbb{1}_{\delta(p) \neq \delta(q)}$,

$$E_{data}(\delta, p) = \begin{cases} e^{-\frac{1}{2\sigma^2}(L(p)-L_S)^2} & \text{if } \delta(p) = 0 \\ e^{-\frac{1}{2\sigma^2}(L(p)-L_L)^2} & \text{if } \delta(p) = 1 \end{cases}$$

et γ est un terme de pondération.

- Minimisation de l'énergie par Graph-Cut.

Correction des Ombres

La luminance d'un point d'une surface Lambertienne est définie par:

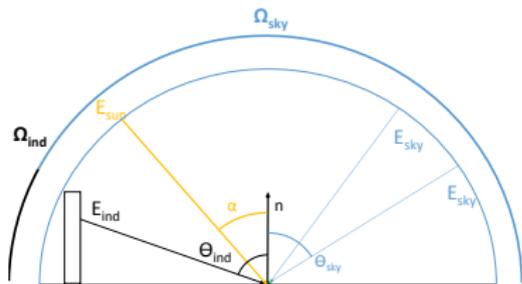
$$L = R(S_{sun} + S_{sky} + S_{indirect})$$

$$S_{sun} = V_{sun} \cos \alpha E_{sun}$$

$$S_{sky} = \int_{\Omega_{sky}} \cos \theta_{sky} E_{sky}$$

$$S_{indirect} = \int_{\Omega_{indirect}} \cos \theta_{indirect} E_{indirect}$$

où I est l'illumination, R l'albedo du matériau, S_{sun} l'éclairage dû au soleil, S_{sky} l'éclairage dû au ciel et $S_{indirect}$ l'éclairage indirect.



Correction des ombres

- Pour les points d'ombre, la contribution du soleil est nulle
- $S_{sky} + S_{indirect} \propto \beta$ est approximé en regardant la différence entre la normale au sol et la normale du point $\beta(\mathbf{p}) = \langle \mathbf{n}(\mathbf{p}), \mathbf{n}_{ground} \rangle$
- Correction de la luminance des points d'ombre:

$$L'(\mathbf{p}) = L(\mathbf{p}) \frac{\bar{\beta}}{\beta(\mathbf{p})}$$

Colorisation finale du nuage de points



(a) Original cloud color.



(b) Reflectance values.



(c) Shadow mask.



(d) Relighted cloud (light/dark).

Colorisation finale du nuage de points



(a) Original Cloud.



(b) Shadow mask.



(c) Relighted cloud (light).



(d) Relighted cloud (dark).

Conclusion

- Vers une collaboration des différents types de données
- Techniques images efficaces qu'on peut transférer au nuage de points
- Pour aller plus loin: densification du nuage de points
- Et encore plus loin: correction des images (retrait des piétons, des voitures...)

References

- *Fine Scale Image Registration in Large-Scale Urban LIDAR Point Sets*, Maximilien Guislain, Julie Digne, Raphaëlle Chaine, Gilles Monnier, Computer Vision and Image Understanding (CVIU) Special Issue on Large-Scale 3D Modeling of Urban Indoor or Outdoor Scenes from Images and Range Scans, Volume 157, pages 90-102, issn 1077-3142, 2017
- *Detecting and Correcting Shadows in Urban Point Clouds and Image Collections*, Maximilien Guislain, Julie Digne, Raphaëlle Chaine, Dimitri Kudelski, Pascal Lefebvre-Albaret, Proceedings 3DV2016, Stanford, October 2016.