

Soutenance de thèse

DÉTECTION DE CHANGEMENTS ENTRE VIDÉOS AÉRIENNES AVEC TRAJECTOIRES ARBITRAIRES

Nicolas BOURDIS

(nicolas.bourdis.2008@ingenieurs-supelec.org)

Directeur de thèse: **Hichem SAHBI**

Co-encadrement de la thèse: **Denis MARRAUD**

24 mai 2013



Contexte opérationnel de l'analyse vidéo

Durant les 5 dernières années :

- Démocratisation des applications de l'analyse vidéo (e.g. triplement du nombre de caméras de vidéo-surveillance autorisées en France entre 2007 et 2012, source CNIL)



Opérateurs de vidéo-surveillance
(Commune de Nice)



Pilote et opérateur drone
(US Air Force)

Durant les 5 dernières années :

- Démocratisation des applications de l'analyse vidéo (e.g. triplement du nombre de caméras de vidéo-surveillance autorisées en France entre 2007 et 2012, source CNIL)
- Progrès technologique significatif vis-à-vis des capacités d'acquisition (i.e. fréquence, taille d'image, capacité de stockage etc)

Système ARGUS (DARPA)

Résolution au sol de 15cm sur 25km^2

Acquisition 24h/24 à 12 image/s

6000 *To* de données générées par jour



Durant les 5 dernières années :

- Démocratisation des applications de l'analyse vidéo (e.g. triplement du nombre de caméras de vidéo-surveillance autorisées en France entre 2007 et 2012, source CNIL)
- Progrès technologique significatif vis-à-vis des capacités d'acquisition (i.e. fréquence, taille d'image, capacité de stockage etc)

⇒ **Explosion du volume de données vidéo à traiter**

Approche semi-automatique de détection de changements

Difficultés :

- Incohérence visuelle (bruit, gamme dynamique, artefacts optique/compression, flou, etc)
- Variabilité des points de vue
- Conditions d'illumination très diverses

(Vidéos aériennes fournies par Cassidian dans le cadre du projet ASIP 2)

- 1 Introduction
 - Problématique
 - Approche proposée
- 2 Prétraitement des données
 - Géo-localisation
 - Illumination
- 3 Critères de détection de changements
 - Base de données 3D
 - Modélisation d'apparence
- 4 Consolidation des résultats
 - Consolidation temporelle
 - Binarisation
 - Retour interactif
- 5 Conclusion

- 1 Introduction
 - Problématique
 - Approche proposée
- 2 Prétraitement des données
- 3 Critères de détection de changements
- 4 Consolidation des résultats
- 5 Conclusion

Définition du problème de la détection de changements

Définition

Détection de changements = Tâche de comparaison de données d'observation **radiométriques** acquises à des dates **différentes**, visant à déterminer la présence et la localisation de changements **significatifs**.

Définition du problème de la détection de changements

Définition

Détection de changements = Tâche de comparaison de données d'observation **radiométriques** acquises à des dates **différentes**, visant à déterminer la présence et la localisation de changements **significatifs**.

Exemple simplifié : Jeu des 7 erreurs

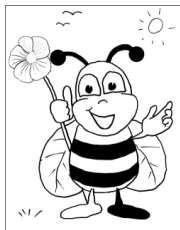


Image 1



Solution

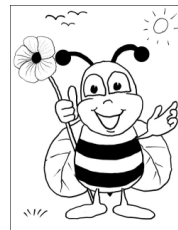


Image 2

Définition du problème de la détection de changements

Définition

Détection de changements = Tâche de comparaison de données d'observation **radiométriques** acquises à des dates **différentes**, visant à déterminer la présence et la localisation de changements **significatifs**.

Cas général



Donnée
de référence



Masque de
changement

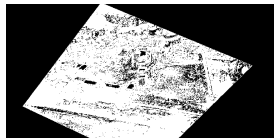


Donnée
de test

Variations fréquentes et non-pertinentes



Référence



Différence seuillée

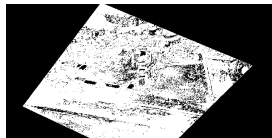


Test

Variations fréquentes et non-pertinentes



Référence



Différence seuillée



Test

Contenu de la scène

- Illumination
- Ombres
- Mouvements répétitifs
- ...

Processus d'acquisition

- Points de vue
- Bruit
- Artefacts
- ...

Changements significatifs

Définition dépendante de l'application visée :

- Conteneurs, objets enterrés, ... pour la détection d'EEI,
- Personnes, vêtements, objets personnels, ... pour la recherche de survivants,
- Feux, personnes, véhicules, ... pour la surveillance de ressources naturelles,
- Catégories de sol ou de végétation pour le suivi de culture,
- ...

Changements significatifs

Définition dépendante de l'application visée :

- Conteneurs, objets enterrés, ... pour la détection d'EEI,
- Personnes, vêtements, objets personnels, ... pour la recherche de survivants,
- Feux, personnes, véhicules, ... pour la surveillance de ressources naturelles,
- Catégories de sol ou de végétation pour le suivi de culture,
- ...

Définition utilisée (scénario opérationnel visé)

- Personnes ou objets artificiels
- Fixes dans la scène

⇒ Exploitation d'information a priori pour affiner les résultats

Évaluation quantitative

Critères d'évaluation

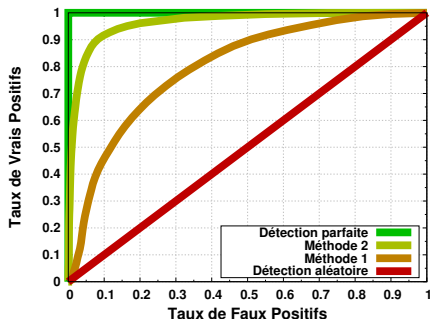
- Taux de vrais positifs (ou rappel) : $p(\text{détection}|\text{changement})$
- Taux de faux positifs : $p(\text{détection}|\overline{\text{changement}})$
- Précision : $p(\text{changement}|\text{détection})$

Évaluation quantitative

Critères d'évaluation

- Taux de vrais positifs (ou rappel) : $p(\text{détection}|\text{changement})$
- Taux de faux positifs : $p(\text{détection}|\overline{\text{changement}})$
- Précision : $p(\text{changement}|\text{détection})$

Courbe ROC



Motivations

Exploitation de la redondance

- Exploitation de la totalité des données disponibles (référence)
- Possibilité de fusion hors-ligne des données de référence
⇒ Accélération du traitement des données de test
- Caractérisation plus précise des changements survenus (test)

Motivations

Exploitation de la redondance

...

Modélisation incrémentale asymétrique

- Énorme volume d'observations
⇒ Compression incrémentale
- Changements pertinents rares et très divers
⇒ Apprentissage asymétrique

Motivations

Exploitation de
la redondance

Modélisation
incrémentale

Représentation tridimensionnelle

- Gestion simple et précise des effets géométriques
- Possibilité de modélisation en coordonnées réelles
⇒ Exploitation jointe avec d'autres types de données

Motivations

Exploitation de
la redondance

Modélisation
incrémentale

Représentation
tridimensionnelle

Approche semi-automatique

- Automatisation des tâches fastidieuses et répétitives
- Bonne robustesse dans les cas ambigus

Motivations

Exploitation de
la redondance

...

Modélisation
incrémentale

...

Représentation
tridimensionnelle

...

Approche semi-
automatique

...

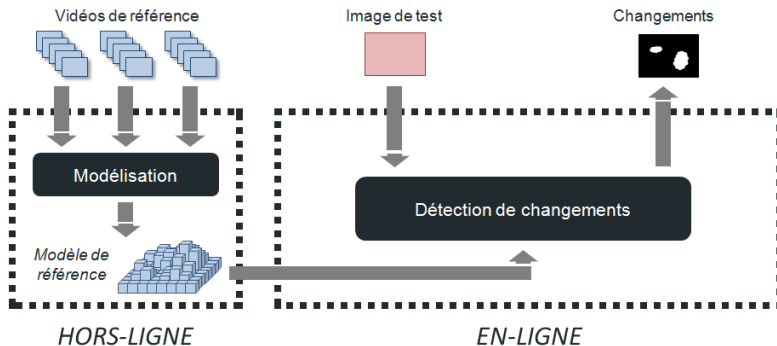
Gestion des variations d'illumination

Combinaison de plusieurs techniques pour réduire les fausses alarmes et non-détections au minimum.

Approches existantes

Approche	Géo-localisation	Effets géométriques	Illumination	Redondance (ref)	Redondance (test)	Semi-automatique	En ligne
Mittal (2000)		·		✓			✓
Primdahl (2005)	·	·	✓				✓
Pollard (2008)		✓	✓	✓			✓
Buchanan (2009)	✓	✓					
Stennet (2009)		·					
Crispell (2012)	✓	✓	·	✓			✓
Approche proposée	✓	✓	✓	✓	✓	✓	✓

Vue d'ensemble de l'approche proposée



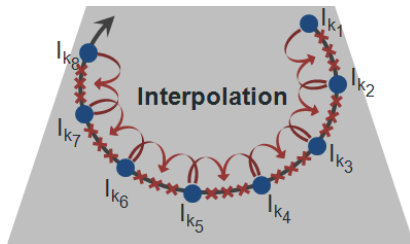
- 1 Introduction
- 2 Prétraitement des données
 - Géolocalisation
 - Illumination
- 3 Critères de détection de changements
- 4 Consolidation des résultats
- 5 Conclusion

Interpolation de poses

Principe et application

Principe

- Calibration manuelle de quelques images-clés
- Propagation automatique aux images restantes

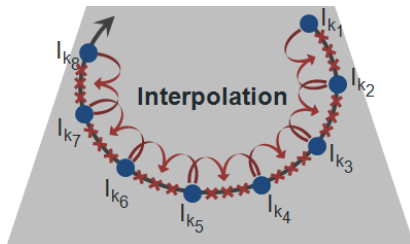


Interpolation de poses

Principe et application

Principe

- Calibration manuelle de quelques images-clés
- Propagation automatique aux images restantes



Utilisation

Approche appropriée pour les vidéos de référence :

- Nécessite l'ensemble des images de la vidéo
- Calibration manuelle des images-clés

Interpolation de poses

Calibration des images-clés

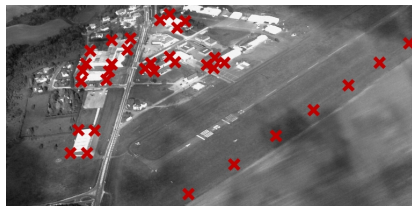
- 1 Définition des images-clés
 $\{I_k\}_{k \in K}$, $K \subset \mathbb{N}$
- 2 Définition d'amers physiques
 $\{M_i\}_{i \in \mathbb{N}}$



Interpolation de poses

Calibration des images-clés

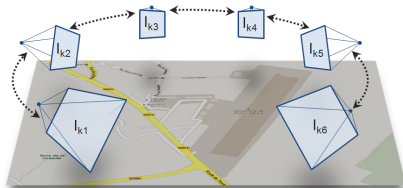
- 1 Définition des images-clés
 $\{I_k\}_{k \in K}$, $K \subset \mathbb{N}$
- 2 Définition d'amers physiques
 $\{\mathbf{M}_i\}_{i \in \mathbb{N}}$
- 3 Saisie des projections
 $\{\mathbf{m}_{i,k}\}_{i \in \mathbb{N}, k \in K}$ de ces amers



Interpolation de poses

Calibration des images-clés

- 1 Définition des images-clés
 $\{I_k\}_{k \in K}$, $K \subset \mathbb{N}$
- 2 Définition d'amers physiques
 $\{M_i\}_{i \in \mathbb{N}}$
- 3 Saisie des projections
 $\{m_{i,k}\}_{i \in \mathbb{N}, k \in K}$ de ces amers
- 4 Estimation ^{1, Alg 7.1} des matrices de projection
 $\{\widehat{P}_k\}_{k \in K}$
- 5 Ajustements de faisceaux
^{1, § 18.1}



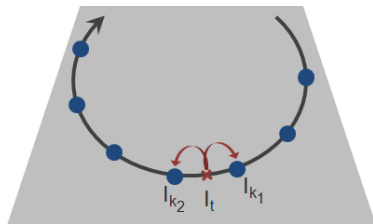
1. Hartley, R. et A. Zisserman. 2000, *Multiple view geometry*, 2^e éd., Cambridge university press

Interpolation de poses

Propagation aux autres images

Pour chaque image intermédiaire $\{I_t\}_{t \in \mathbb{N} \setminus K}$:

- 1 Identification des images-clés voisines I_{k_1} et I_{k_2} telles que $k_1 = \sup \{k \in K \mid k < t\}$ et $k_2 = \inf \{k \in K \mid k > t\}$

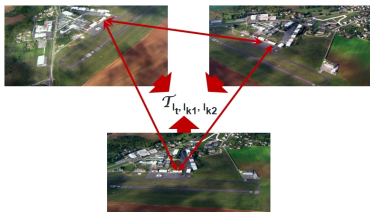


Interpolation de poses

Propagation aux autres images

Pour chaque image intermédiaire $\{I_t\}_{t \in \mathbb{N} \setminus K}$:

- 1 Identification des images-clés voisines I_{k_1} et I_{k_2} telles que $k_1 = \sup \{k \in K \mid k < t\}$ et $k_2 = \inf \{k \in K \mid k > t\}$
- 2 Calcul des correspondances (>6) entre I_t , I_{k_1} et I_{k_2}
- 3 Estimation ^{1, algorithme 16.4} du tenseur trifocal $\hat{\mathcal{T}}_{t,k_1,k_2}$



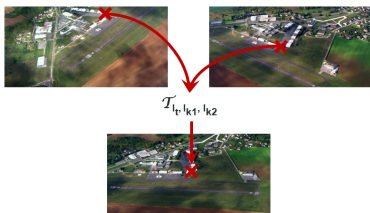
1. Hartley, R. et A. Zisserman. 2000, *Multiple view geometry*, 2^e éd., Cambridge university press

Interpolation de poses

Propagation aux autres images

Pour chaque image intermédiaire $\{I_t\}_{t \in \mathbb{N} \setminus K}$:

- 1 Identification des images-clés voisines I_{k_1} et I_{k_2} telles que $k_1 = \sup \{k \in K \mid k < t\}$ et $k_2 = \inf \{k \in K \mid k > t\}$
- 2 Calcul des correspondances (>6) entre I_t , I_{k_1} et I_{k_2}
- 3 Estimation ^{1, algorithme 16.4} du tenseur trifocal $\hat{\mathcal{T}}_{t,k_1,k_2}$
- 4 Estimation ^{1, § 15.3.2} des projections $\{\hat{\mathbf{m}}_{i,t}\}_{i \in \mathbb{N}}$
- 5 Estimation ^{1, Alg 7.1} de la matrice de projection $\{\hat{\mathbf{P}}_t\}_{k \in K}$



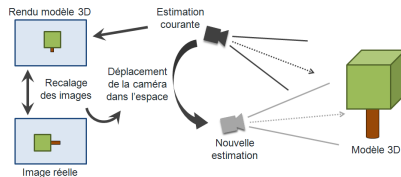
1. Hartley, R. et A. Zisserman. 2000, *Multiple view geometry*, 2^e éd., Cambridge university press

Asservissement visuel

Principe et application

Principe

- Rendu du modèle 3D depuis l'estimation courante
- Ajustement pour aligner le rendu et l'image réelle ¹



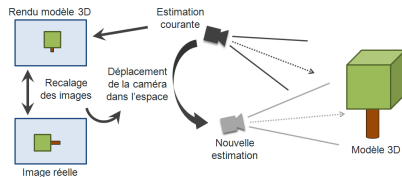
1. Crispell, D., J. Mundy et G. Taubin. 2008, « Parallax-free registration of aerial video », dans *Proceedings of the British Machine Vision Conference*, p. 73.1–73.10

Asservissement visuel

Principe et application

Principe

- Rendu du modèle 3D depuis l'estimation courante
- Ajustement pour aligner le rendu et l'image réelle ¹



Utilisation

Approche appropriée pour les vidéos de test :

- Traitement en ligne d'une vidéo
- Nécessite modèle 3D de la scène

1. Crispell, D., J. Mundy et G. Taubin. 2008, « Parallax-free registration of aerial video », dans *Proceedings of the British Machine Vision Conference*, p. 73.1–73.10

Asservissement visuel

Étape de prédiction

Objectif

Prédiction des paramètres d'acquisition courants, à l'aide du recalage observé entre les images réelles précédente et courante.

Asservissement visuel

Étape de prédiction

Objectif

Prédiction des paramètres d'acquisition courants, à l'aide du recalage observé entre les images réelles précédente et courante.

- 1 Estimation du recalage $\hat{H}_{n \leftarrow n-1}$ entre les images précédente et courante
- 2 Recherche (Méthode de Newton) d'une solution $\tilde{\mathbf{x}}_n$ de l'équation

$$F(\mathbf{x}) = \underbrace{\text{vec} \left(H_{n \leftarrow n-1}^f(\mathbf{x}_{n-1}, \mathbf{x}, \mathbf{n}_\pi, d_\pi) \right)}_{\text{Recalage analytique}} - \underbrace{\text{vec} \left(\hat{H}_{n \leftarrow n-1} \right)}_{\text{Recalage mesuré}} = 0$$

Asservissement visuel

Étape de prédiction

Objectif

Prédiction des paramètres d'acquisition courants, à l'aide du recalage observé entre les images réelles précédente et courante.

- 1 Estimation du recalage $\hat{H}_{n \leftarrow n-1}$ entre les images précédente et courante
- 2 Recherche (Méthode de Newton) d'une solution $\tilde{\mathbf{x}}_n$ de l'équation

$$F(\mathbf{x}) = \underbrace{\text{vec} \left(H_{n \leftarrow n-1}^f(\mathbf{x}_{n-1}, \mathbf{x}, \mathbf{n}_\pi, d_\pi) \right)}_{\text{Recalage analytique}} - \underbrace{\text{vec} \left(\hat{H}_{n \leftarrow n-1} \right)}_{\text{Recalage mesuré}} = 0$$

Problème : Accumulation d'erreur \Rightarrow divergence de l'estimation

Asservissement visuel

Étape de correction

Objectif

Correction de l'estimation des paramètres d'acquisition à l'aide d'une référence absolue (modèle 3D).

Asservissement visuel

Étape de correction

Objectif

Correction de l'estimation des paramètres d'acquisition à l'aide d'une référence absolue (modèle 3D).

Itérer jusqu'à convergence :

- 1 Rendu $I_r(\tilde{\mathbf{x}}_n)$ du modèle 3D selon l'estimation courante des paramètres d'acquisition $\tilde{\mathbf{x}}_n$
- 2 Recherche (Méthode de Newton) d'une solution $\hat{\mathbf{x}}_n$ de l'équation $F(\mathbf{x}) = \text{vec}(\underbrace{\hat{H}_{r \leftarrow n}(\mathbf{x})}_{\text{Recalage mesuré}}) - \text{vec}(\text{ID}) = 0$

Asservissement visuel

Étape de correction

Objectif

Correction de l'estimation des paramètres d'acquisition à l'aide d'une référence absolue (modèle 3D).

Itérer jusqu'à convergence :

- 1 Rendu $I_r(\tilde{\mathbf{x}}_n)$ du modèle 3D selon l'estimation courante des paramètres d'acquisition $\tilde{\mathbf{x}}_n$
- 2 Recherche (Méthode de Newton) d'une solution $\hat{\mathbf{x}}_n$ de l'équation $F(\mathbf{x}) = \text{vec}(\underbrace{\hat{H}_{r \leftarrow n}(\mathbf{x})}_{\text{Recalage mesuré}}) - \text{vec}(\text{ID}) = 0$

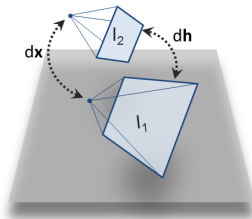
Variante : ré-utiliser l'étape de prédiction sur I_n et $I_r(\tilde{\mathbf{x}}_n)$

Asservissement visuel

Matrice Jacobienne

Rôle

Lie une modification infinitésimale des paramètres d'acquisition à la modification infinitésimale de la transformation de recalage correspondante : $J \, dx = dh$

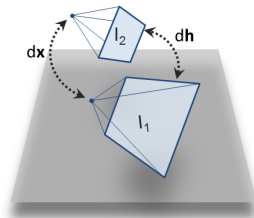


Asservissement visuel

Matrice Jacobienne

Rôle

Lie une modification infinitésimale des paramètres d'acquisition à la modification infinitésimale de la transformation de recalage correspondante : $J \, dx = dh$



Cas restreint

Paramètres extrinsèques

$$J \in \mathcal{M}_{8 \times 6}(\mathbb{R})$$

Cas général

Paramètres extrinsèques
et intrinsèques

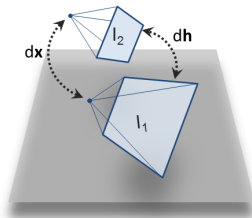
$$J \in \mathcal{M}_{8 \times 10}(\mathbb{R})$$

Asservissement visuel

Matrice Jacobienne

Rôle

Lie une modification infinitésimale des paramètres d'acquisition à la modification infinitésimale de la transformation de recalage correspondante : $J dx = dh$



Cas restreint

Paramètres extrinsèques

$$J \in \mathcal{M}_{8 \times 6}(\mathbb{R})$$

Système sur-contraint

Cas général

Paramètres extrinsèques
et intrinsèques

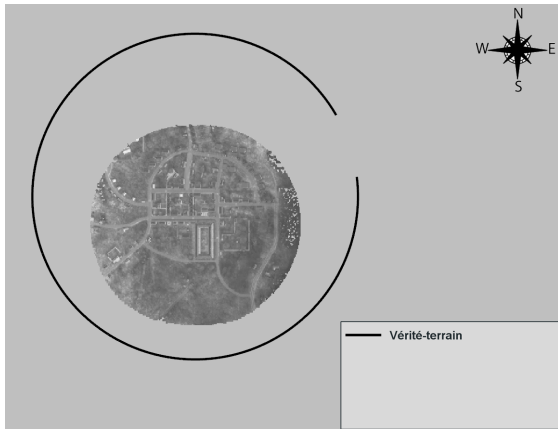
$$J \in \mathcal{M}_{8 \times 10}(\mathbb{R})$$

Système sous-contraint

Comparaison des techniques de géo-localisation

Erreur sur les paramètres d'acquisition

Exemple de trajectoires estimées sur données synthétiques



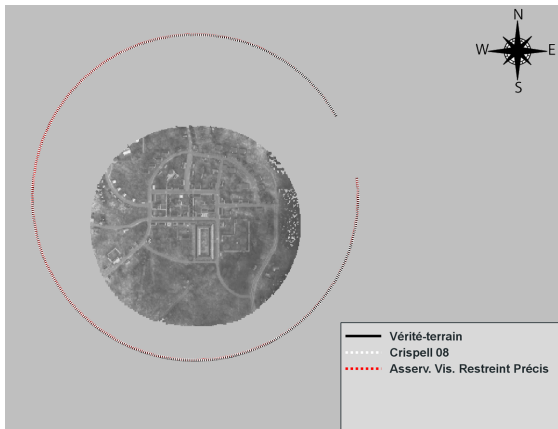
Temps d'exécution par image

Méthode	Temps (s)

Comparaison des techniques de géo-localisation

Erreur sur les paramètres d'acquisition

Exemple de trajectoires estimées sur données synthétiques



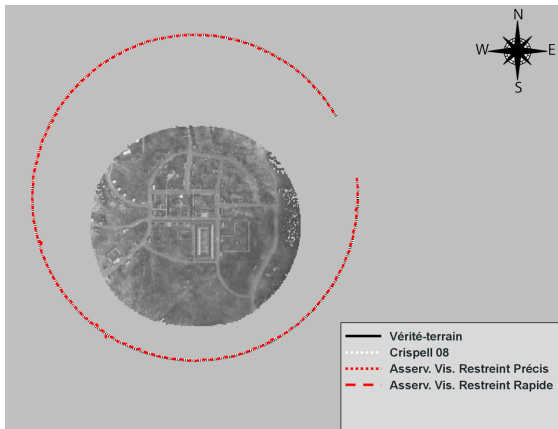
Temps d'exécution par image

Méthode	Temps (s)
Crispell 08	26.6
Asserv. Visuel Restreint Précis	23.0

Comparaison des techniques de géo-localisation

Erreur sur les paramètres d'acquisition

Exemple de trajectoires estimées sur données synthétiques



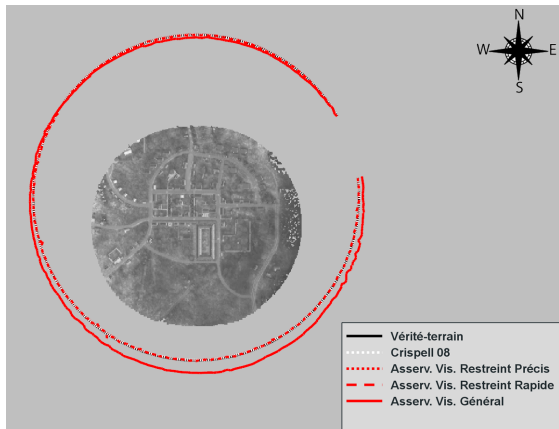
Temps d'exécution par image

Méthode	Temps (s)
Crispell 08	26.6
Asserv. Visuel Restreint Précis	23.0
Asserv. Visuel Restreint Rapide	5.0

Comparaison des techniques de géo-localisation

Erreur sur les paramètres d'acquisition

Exemple de trajectoires estimées sur données synthétiques



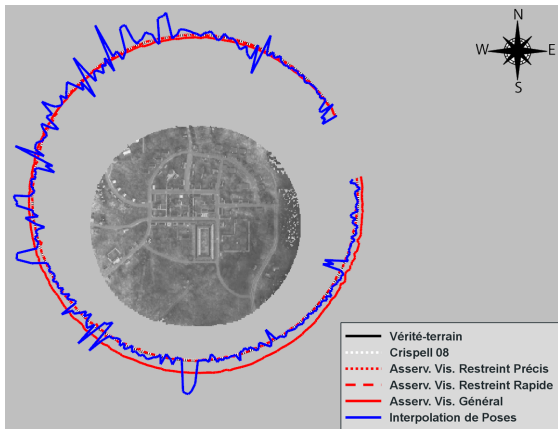
Temps d'exécution par image

Méthode	Temps (s)
Crispell 08	26.6
Asserv. Visuel Restreint Précis	23.0
Asserv. Visuel Restreint Rapide	5.0
Asserv. Visuel Général	25.0

Comparaison des techniques de géo-localisation

Erreur sur les paramètres d'acquisition

Exemple de trajectoires estimées sur données synthétiques



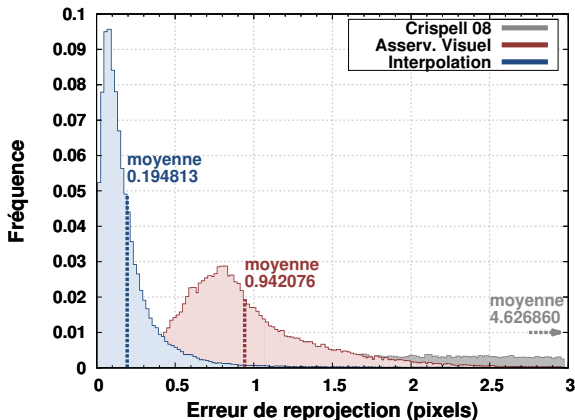
Temps d'exécution par image

Méthode	Temps (s)
Crispell 08	26.6
Asserv. Visuel Restreint Précis	23.0
Asserv. Visuel Restreint Rapide	5.0
Asserv. Visuel Général	25.0
Interpolation	32.9

Comparaison des techniques de géo-localisation

Erreur de recalage image / modèle

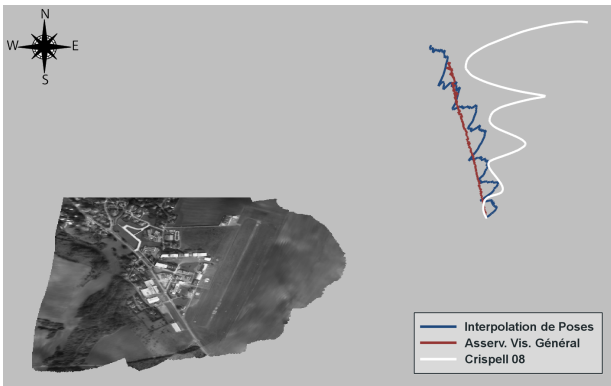
Histogrammes des erreurs de reprojection



Comparaison des techniques de géo-localisation

Données réelles

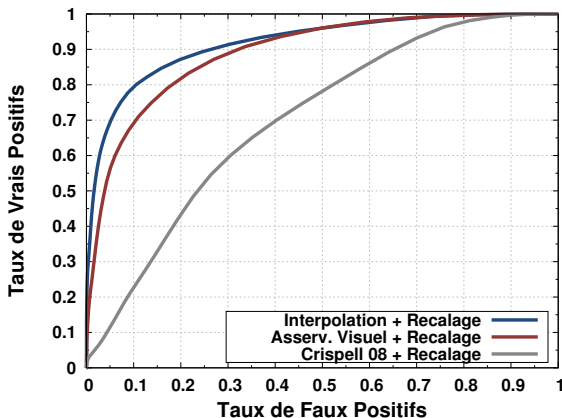
Trajectoires estimées sur données réelles



Comparaison des techniques de géo-localisation

Données réelles

Influence sur les performances en détection de changements



Problème et approche

Variations d'illumination \Leftrightarrow **Variations visuelles considérables**
(e.g. zones éclairées, ombres portées)

Premier mécanisme

Atténuation des effets de l'illumination préalablement à la détection de changements.

Problème et approche

Variations d'illumination \Leftrightarrow **Variations visuelles considérables**
(e.g. zones éclairées, ombres portées)

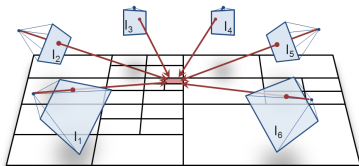
Premier mécanisme

Atténuation des effets de l'illumination préalablement à la détection de changements.

Notion de chromaticité

Permet :

- Traitement indépendant en chaque cellule
- Exploitation des observations de référence multiples



Coordonnées chromatiques en représentation brute

1. Gevers, T. et W. M. Smeulders. 1999, « Color based object recognition », *Pattern recognition*, vol. 32, n° 3, p. 453–464
2. Finlayson, G., S. Hordley, C. Lu et M. Drew. 2006, « On the removal of shadows from images », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, n° 1, p. 59–68

Coordonnées chromatiques en représentation brute

Problème : Perte d'information

1. Gevers, T. et W. M. Smeulders. 1999, « Color based object recognition », *Pattern recognition*, vol. 32, n° 3, p. 453–464
2. Finlayson, G., S. Hordley, C. Lu et M. Drew. 2006, « On the removal of shadows from images », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, n° 1, p. 59–68

Représentations transformées (CC. classiques)

Représentation normalisée

$$\begin{pmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{pmatrix} = \begin{pmatrix} \frac{R}{R+G+B} \\ \frac{G}{R+G+B} \\ \frac{B}{R+G+B} \end{pmatrix} \cdot (\bar{R}_{ref} + \bar{G}_{ref} + \bar{B}_{ref})$$

Représentation compensée

$$\begin{pmatrix} R_{comp} \\ G_{comp} \\ B_{comp} \end{pmatrix} = (1 - \beta) \cdot \begin{pmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{pmatrix} + \beta \cdot \begin{pmatrix} \bar{R}_{ref} \\ \bar{G}_{ref} \\ \bar{B}_{ref} \end{pmatrix}$$

$$\text{avec } \beta = \left(1 - \frac{R + G + B}{3}\right)^5$$

Représentations transformées (CC. classiques)

Représentation normalisée

$$\begin{pmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{pmatrix} = \begin{pmatrix} \frac{R}{R+G+B} \\ \frac{G}{R+G+B} \\ \frac{B}{R+G+B} \end{pmatrix} \cdot (\bar{R}_{ref} + \bar{G}_{ref} + \bar{B}_{ref})$$

Peut causer des faux positifs

Représentation compensée

$$\begin{pmatrix} R_{comp} \\ G_{comp} \\ B_{comp} \end{pmatrix} = (1 - \beta) \cdot \begin{pmatrix} R_{norm} \\ G_{norm} \\ B_{norm} \end{pmatrix} + \beta \cdot \begin{pmatrix} \bar{R}_{ref} \\ \bar{G}_{ref} \\ \bar{B}_{ref} \end{pmatrix}$$

$$\text{avec } \beta = \left(1 - \frac{R + G + B}{3}\right)^5$$

Peut causer des non-détections

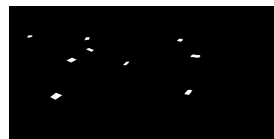
Exemples de masques de changements estimés



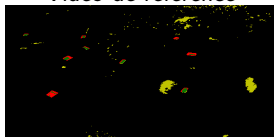
Vidéo de référence



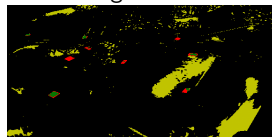
Image de test



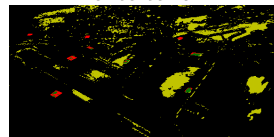
Vérité-terrain



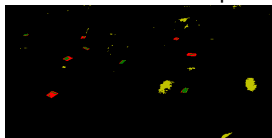
Rep. brute CC. class.



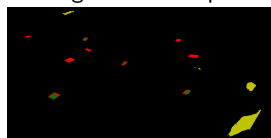
Rep. brute CC. log.



Rep. brute CC. L1L2L3



Rep. comp. CC. class.

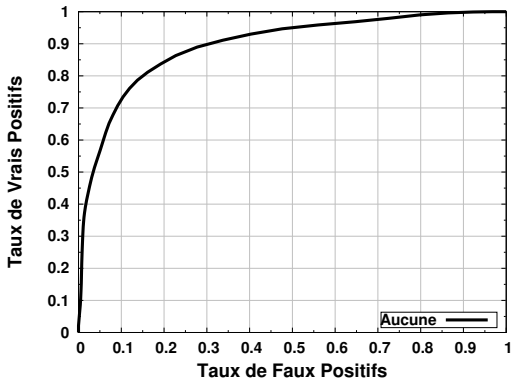


Sans traitement

Vrais positifs
 Faux positifs
 Faux négatifs

Comparaison des représentations

Courbes ROC en fonction de la technique de représentation

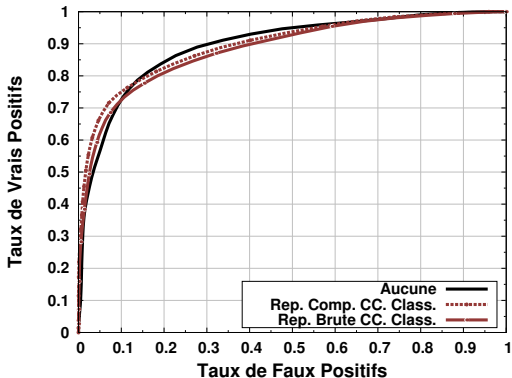


Temps d'exécution par image

Méthode	Temps (ms)
Aucune	30.5

Comparaison des représentations

Courbes ROC en fonction de la technique de représentation

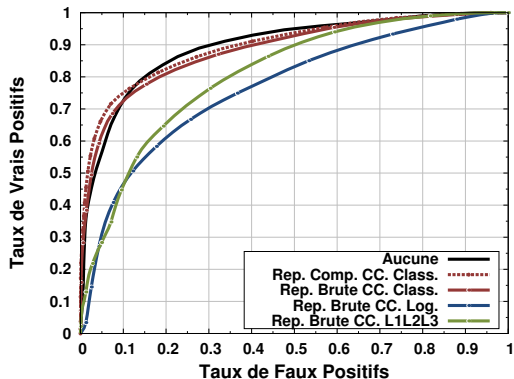


Temps d'exécution par image

Méthode	Temps (ms)
Aucune	30.5
Rep. comp. CC. class.	37.7
Rep. brute CC. class.	35.2

Comparaison des représentations

Courbes ROC en fonction de la technique de représentation



Temps d'exécution par image

Méthode	Temps (ms)
Aucune	30.5
Rep. comp. CC. class.	37.7
Rep. brute CC. class.	35.2
Rep. brute CC. log.	128.9
Rep. brute CC. L1L2L3	36.9

- 1 Introduction
- 2 Prétraitement des données
- 3 Critères de détection de changements**
 - Base de données 3D
 - Modélisation d'apparence
- 4 Consolidation des résultats
- 5 Conclusion

Organisation des modèles d'apparence

Structure utilisée

Quad-Tree augmenté

- Quad-Tree – subdivision bi-dimensionnelle à résolution adaptative
- Carte d'élévation – altitude associée à chaque cellule

Base de données 3D

Organisation des modèles d'apparence

Méthode de génération de la base de données

Entrées nécessaires

- Modèle tridimensionnel (e.g. MNT, MNE, reconstruction 3D)
- Paramètres d'acquisition pour chaque image

Organisation des modèles d'apparence

Méthode de génération de la base de données

Entrées nécessaires

- Modèle tridimensionnel (e.g. MNT, MNE, reconstruction 3D)
- Paramètres d'acquisition pour chaque image

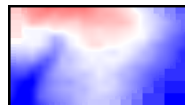
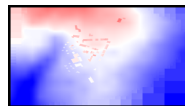
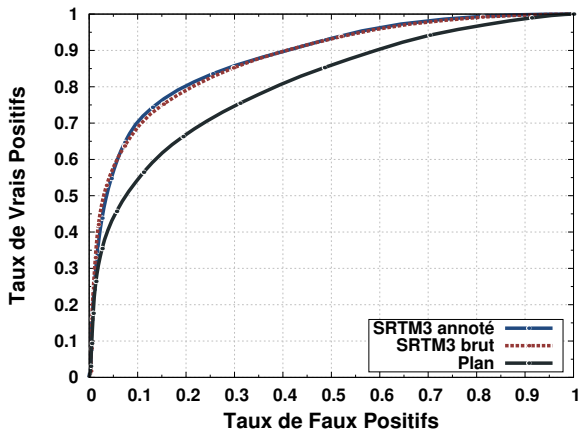
Étapes de la génération

- 1 Définition de la racine du Quad-Tree
- 2 Adaptation de la résolution aux observations de référence
- 3 Calcul de l'élévation de chaque cellule



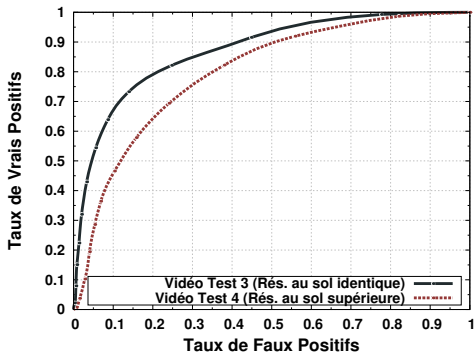
Influence de la précision des élévations

Courbes ROC en fonction de la précision des élévations



Influence de l'écart de résolution au sol

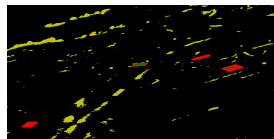
Courbes ROC en fonction de l'écart de résolution au sol



Vidéo de référence



Image de test (vidéo 4)



Résultat de détection

Rôle de la modélisation d'apparence

Objectifs

- Compression incrémentale des observations de référence
- Modélisation des variations normales dans la scène
- Comparaison d'une observation de test à un modèle de référence

Rôle de la modélisation d'apparence

Objectifs

- Compression incrémentale des observations de référence
- Modélisation des variations normales dans la scène
- Comparaison d'une observation de test à un modèle de référence

Critères de comparaison des techniques

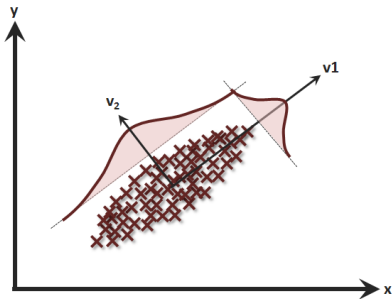
- Exploitation du contexte (nul, local ou global)
- Modélisation uni-modale ou multi-modale

Modélisation par ACP incrémentale

Principe

Caractéristique

Modélisation multi-modale et exploitant le contexte global



Analyse en composantes principales

Variations modélisées

- Inversions de contrastes,
- Variations d'éclairage,
- Recalage approximatif,
- Apparence multi-modale,
- ...

Modélisation d'apparence

Modélisation par ACP incrémentale

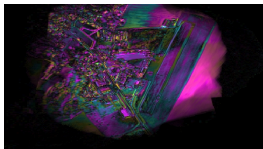
Principe

Caractéristique

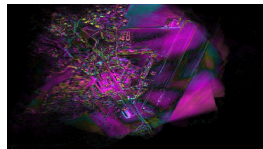
Modélisation multi-modale et exploitant le contexte global



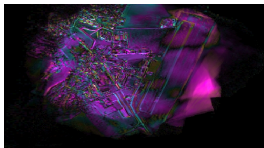
Moyenne



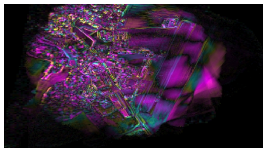
CP 1



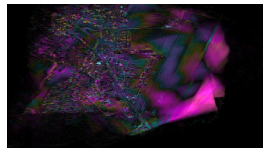
CP 2



CP 3



CP 4

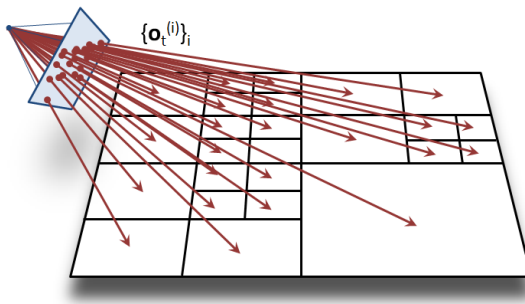


CP 5

Modélisation par ACP incrémentale

Vecteur d'observations

Construction du vecteur d'observations

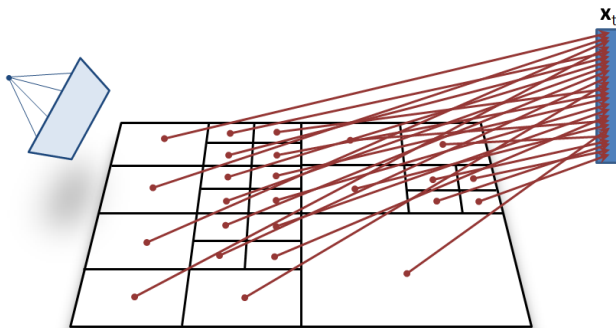


1. Projection des observations dans le modèle 3D

Modélisation par ACP incrémentale

Vecteur d'observations

Construction du vecteur d'observations



2. Ordonnancement des observations dans le vecteur \mathbf{x}_t
(e.g. $\dim(\mathbf{x}_t) = 3 \times 142\,000$)

Modélisation par ACP incrémentale

Adaptation dans le cadre de la détection de changements

Spécificités p/r soustraction de fond

- 1 Comparaison p/r ensemble des données de référence
⇒ Utilisation d'un poids de mise-à-jour dégressif
- 2 Données manquantes fréquentes
(dus aux occultations 3D et vues partielles)
⇒ Nécessité d'usage d'un mécanisme robuste
- 3 Modélisation hors-ligne des données de référence
⇒ Ordre aléatoire de parcours des données

(Formulation initiale ¹ dans le cadre de la soustraction de fond)

1. Li, Y. 2004, « On incremental and robust subspace learning », *Pattern recognition*, vol. 37, n° 7, p. 1509–1518

Modélisation par ACP incrémentale

Mise à jour incrémentale

Calcul incrémental des moyenne et variance

Pour $t \geq 1$:

$$\begin{cases} \boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} + \boldsymbol{\alpha}_t \circ (\mathbf{x}_t - \boldsymbol{\mu}_{t-1}) \\ \mathbf{v}_t = (1 - \boldsymbol{\alpha}_t) \circ [\mathbf{v}_{t-1} + \boldsymbol{\alpha}_t \circ (\mathbf{x}_t - \boldsymbol{\mu}_{t-1})^{\textcircled{2}}] \end{cases}$$

Avec :

$$\boldsymbol{\alpha}_t = \left\{ \alpha_t^i \right\}_i = \left\{ \frac{\delta_t^i}{\sum_{n=0}^t \delta_n^i} \right\}_i$$

$$\boldsymbol{\mu}_0 = \mathbf{x}_0$$

$$\mathbf{v}_0 = \mathbf{0}_{\dim(\mathbf{x}_0)}$$

Modélisation par ACP incrémentale

Mise à jour incrémentale

Calcul incrémental des composantes principales

Pour $k \leq \min(t, N_{CP})$:

$$\begin{cases} \mathbf{u}_t^k = \mathbf{A}_t \mathbf{u}'_t^k \\ \lambda_t^k = \lambda'^k_t \end{cases}$$

Où $\{\mathbf{u}'_t^k, \lambda'^k_t\}$ sont les vecteurs et valeurs propres de la matrice de Gram $(\mathbf{A}_t^T \mathbf{A}_t)$, avec :

$$\mathbf{A}_t = \left[\mathbf{y}_1 \quad \dots \quad \mathbf{y}_{N_{CP}} \quad \sqrt{\frac{1}{t}} \cdot \check{\mathbf{x}}_t \right]$$

$$\mathbf{y}_k = \sqrt{\frac{t}{t+1}} \lambda_{t-1}^k \cdot \mathbf{u}_{t-1}^k$$

$$\check{\mathbf{x}}_t = \delta_t \circ (\mathbf{x}_t - \boldsymbol{\mu}_t)$$

Modélisation par ACP incrémentale

Détection de changements

Score de détection

$$\mathbf{s}(\mathbf{x}) = \frac{1}{\gamma^2} \cdot [\check{\mathbf{x}} - \mathbf{U}\mathbf{U}^T\check{\mathbf{x}}]^{(2)} \odot \mathbf{v}$$

$$\epsilon_{\text{ipca}}(\mathbf{x}) = 1 - \frac{1}{1 + s_R^i(\mathbf{x})^2 + s_G^i(\mathbf{x})^2 + s_B^i(\mathbf{x})^2}$$

Avec :

$$\mathbf{U} = [\mathbf{u}^1 \quad \dots \quad \mathbf{u}^{N_{\text{CP}}}]$$

$$\check{\mathbf{x}} = \delta \circ (\mathbf{x} - \boldsymbol{\mu})$$

γ = paramètre de normalisation

Algorithmes de modélisation comparés

Algorithme	Nombre de modes	Exploitation du contexte
Gaussienne unique (1G)	1	Non
Mélange de gaussiennes (GMM) ¹	3	Non
ACP par blocs (bPCA)	20	Local
SVD incrémentale (iSVD) ²	20	Global
Méthode retenue (iPCA)	20	Global

1. Stauffer, C. et W. Grimson. 2000, « Learning patterns of activity using real-time tracking », *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, n° 8, p. 747–757

2. Brand, M. 2002, « Incremental singular value decomposition of uncertain data with missing values », dans *Proceedings of the European Conference on Computer Vision*, p. 707–720

Exemples de masques de changements estimés



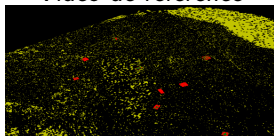
Vidéo de référence



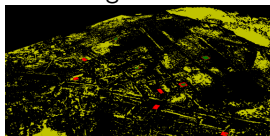
Image de test



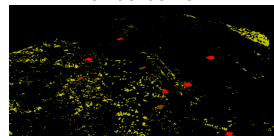
Vérité-terrain



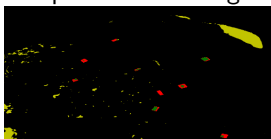
Gaussienne unique



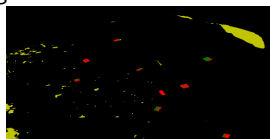
Mélange de gaussiennes



ACP par bloc



SVD incrémentale



ACP incrémentale



Vrais positifs



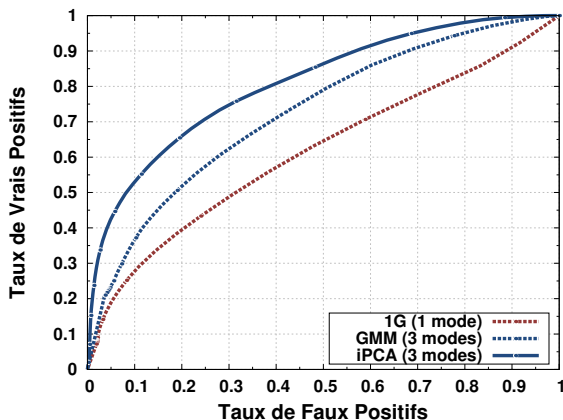
Faux positifs



Faux négatifs

Comparaison des techniques de modélisation d'apparence

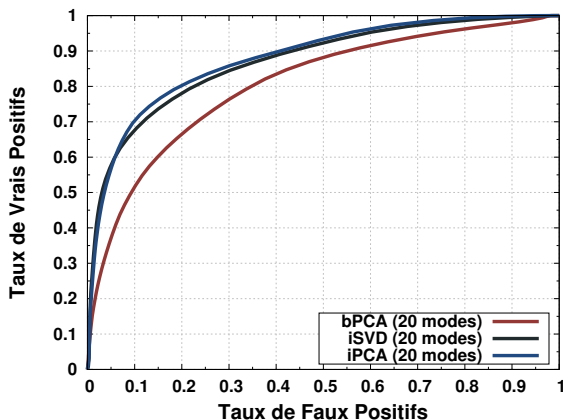
Courbes ROC en fonction de l'algorithme de modélisation



Contexte global vs Aucun contexte

Comparaison des techniques de modélisation d'apparence

Courbes ROC en fonction de l'algorithme de modélisation

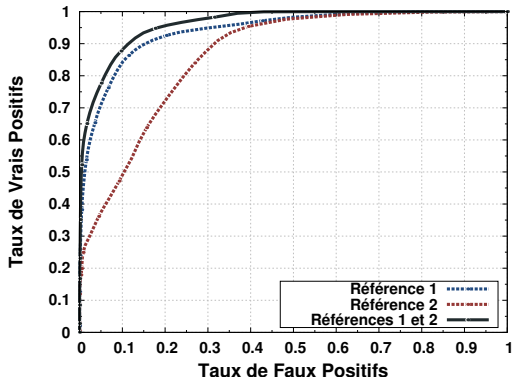


Contexte global vs Contexte local

Analyse de la modélisation d'apparence

Diversité dans les données de référence

Courbes ROC en fonction des données de référence utilisées



Vidéo de référence 1
(Points de vue spécifiques)



Vidéo de référence 2
(Illumination spécifique)



Vidéo de Test

- 1 Introduction
- 2 Prétraitement des données
- 3 Critères de détection de changements
- 4 Consolidation des résultats**
 - Consolidation temporelle
 - Binarisation
 - Retour interactif
- 5 Conclusion

Consolidation temporelle

Optimisation spatio-temporelle

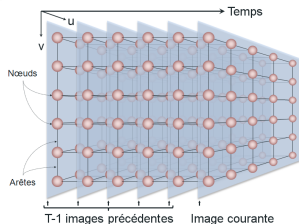
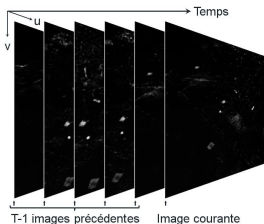
Principe de la propagation de croyance

Objectif

Maximisation de la loi de probabilité jointe :

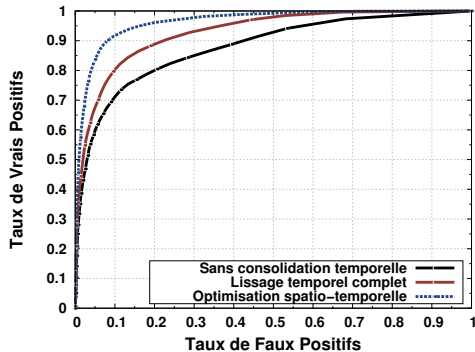
$$p(\{x_l, y_l\}_{l \in \mathcal{L}_t}) = \frac{1}{Z} \exp \left[- E(\{x_l, y_l\}_{l \in \mathcal{L}_t}) \right]$$

$$\text{avec } E(\{x_l, y_l\}_{l \in \mathcal{L}_t}) = - \sum_{n_l \in \mathcal{N}_t} \ln \underbrace{\Phi_{n_l}(y_l, x_l)}_{\text{Attache aux données}} - \sum_{a_{l \leftrightarrow l'} \in \mathcal{A}_t} \ln \underbrace{\Psi_{a_{l \leftrightarrow l'}}(y_l, y_{l'})}_{\text{Fonction d'interaction}}$$



Comparaison des techniques de consolidation temporelle

Courbes ROC en fonction de la technique de consolidation temporelle



Temps d'exécution par image

Méthode	Temps (s)
Lissage temporel	1.6
Optimisation spatio-temporelle	9.9

Binarisation intelligente

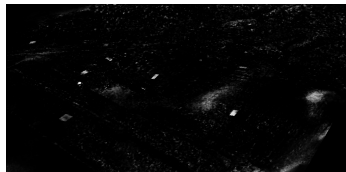
Rôle

Étape de binarisation – Conversion des scores de détection continus en états binaires (*changements* ou *non-changements*)

Binarisation intelligente (MSER¹)

Permet :

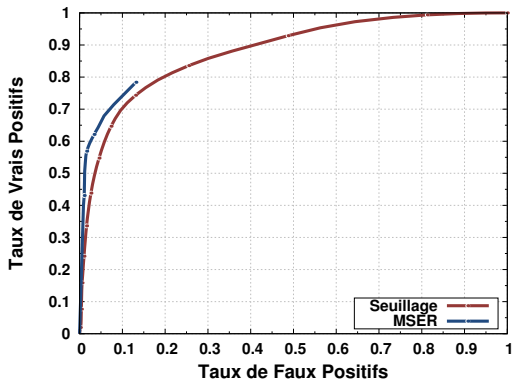
- Extraction de régions stables (retour interactif de pertinence)
- Filtrage des variations de l'illumination résiduelles



1. Matas, J., O. Chum, M. Urban et T. Pajdla. 2004, « Robust wide-baseline stereo from maximally stable extremal regions », *Image and Vision Computing*, vol. 22, n° 10, p. 761–767

Comparaison des techniques de binarisation

Courbes ROC en fonction de la technique de binarisation



Temps
d'exécution par
image

Méthode	Temps (ms)
Seuilage	4
MSEr	417

Principe du mécanisme de retour interactif

Objectif

Exploitation d'information a priori obtenue :

- De manière transparente, par interaction avec l'utilisateur
- De manière flexible, en s'adaptant à l'application désirée
- De manière en ligne, avec la vidéo de test

Principe du mécanisme de retour interactif

Objectif

Exploitation d'information a priori obtenue :

- De manière transparente, par interaction avec l'utilisateur
- De manière flexible, en s'adaptant à l'application désirée
- De manière en ligne, avec la vidéo de test

Approche adoptée

- Traitement des détections fournies par les algorithmes précédents :
 - ⇒ Fonctionnement rapide et transparent
 - ⇒ Réduction des fausses alarmes et conservation des vrais positifs
 - ⇒ Impossibilité de récupération des non-détections
- Classification par SVM linéaire

Descripteur de régions

Exploitation d'un descripteur de région à 22 dimensions

Intensité

Différence d'intensité moyenne,
Valeur majoritaire du rapport
d'intensité

Gradient (bordure de région)

Gradient moyen dans la carte de
scores, Évolution moyenne du
gradient dans les trois canaux
de couleur

Couleur

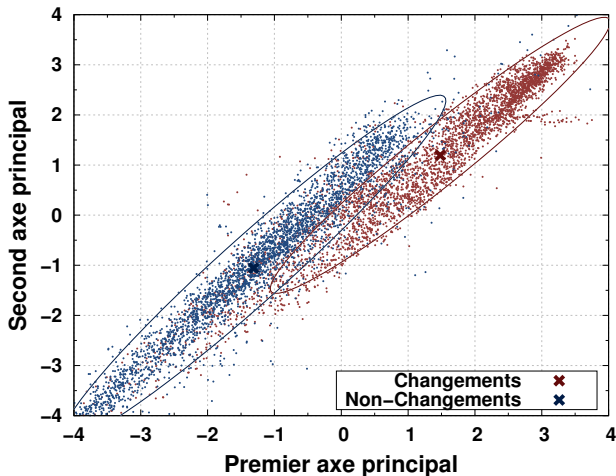
Différence moyenne de teinte,
Valeur majoritaire de proportion
d'information dépendant de
l'illumination

Forme

Taux de variation de l'aire,
Élongation, Non-compacité,
Régularité de bordure

Descripteur de régions

Analyse discriminante

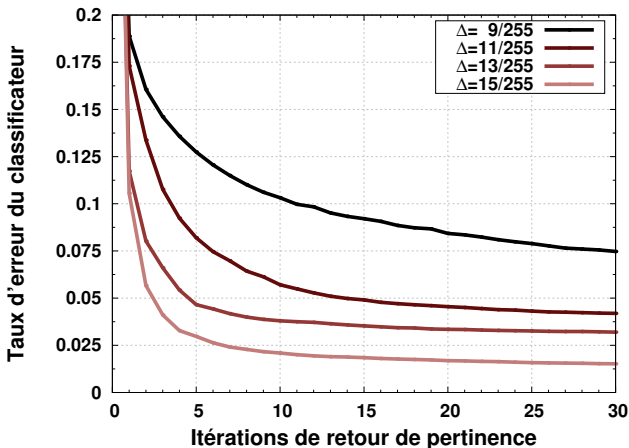


Retour interactif

Performance du retour interactif de pertinence

Performances idéales

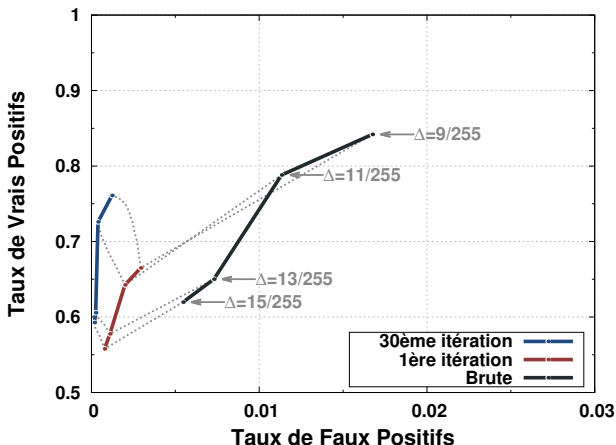
Taux d'erreur du classificateur



Performance du retour interactif de pertinence

Performances idéales

Performances en détection de changements

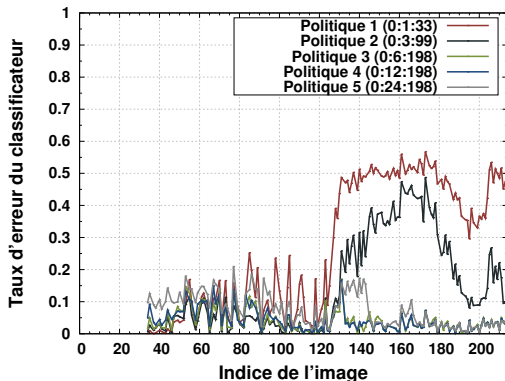


Retour interactif

Performance du retour interactif de pertinence

Performances en ligne

Taux d'erreur selon la politique d'annotation



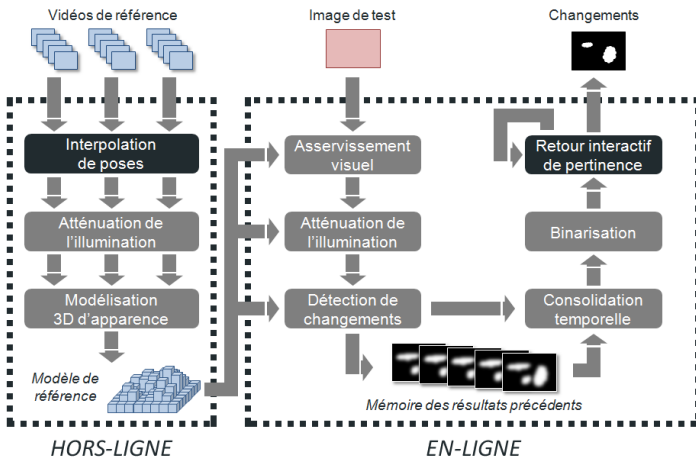
Temps d'exécution par image

Tâche	Temps (ms)
Descripteurs	646
Prédiction	191
Mise à jour	351

- 1 Introduction
- 2 Prétraitement des données
- 3 Critères de détection de changements
- 4 Consolidation des résultats
- 5 Conclusion**

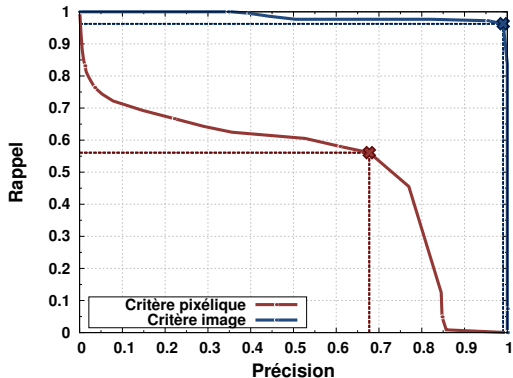
Chaîne de détection de changements complète

Diagramme de fonctionnement global



Performances en Précision / Rappel

Courbes ROC en fonction du critère d'évaluation



Point critère pixelique : (67.7%, 56.1%)

Point critère image : (99.0%, 96.3%)



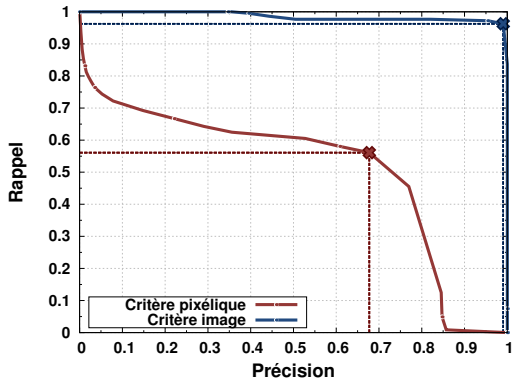
Vidéo de référence



Vidéo de test

Performances en Précision / Rappel

Courbes ROC en fonction du critère d'évaluation



Point critère pixelique : (67.7%, 56.1%)

Point critère image : (99.0%, 96.3%)

⇒ 1 FP / 3 minutes (5 fps)



Vidéo de référence



Vidéo de test

Perspectives

Pistes d'extension des travaux

- **Automatisation et perfectionnement de la géo-localisation**
Désignation automatique d'images-clés, Exploitation conjointe de multiples vidéos, Exploitation des modèles d'apparence, Extension à d'autres types d'application, etc.
- **Approfondissement de la modélisation d'apparence**
Maximisation de la résolution du modèle 3D, Modélisation hiérarchique, Recherche d'algorithmes de modélisation plus performants, etc.
- **Détection de changements hors-ligne**
Comparaison de l'ensemble des observations de référence avec l'ensemble des observations de test, Consolidation par apprentissage actif, etc.
- **Adaptation / portage des algorithmes sur GPU**

Effectué ou accepté

- 1 Bourdis, N., H. Sahbi et D. Marraud. 2011b, « Exploitation de vidéos aériennes multiples pour la détection de changements », cahier de recherche, Telecom ParisTech
- 2 Bourdis, N., D. Marraud et H. Sahbi. 2011a, « Constrained optical flow for aerial image change detection », dans *Proceedings of the IEEE International Geoscience And Remote Sensing Symposium*, IEEE Computer Society, p. 4176–4179
- 3 Bourdis, N. 2011, « Aicd change detection dataset », Base de données hébergée par Computer Vision Online. URL <http://www.computervisiononline.com/dataset/change-detection-dataset>
- 4 Bourdis, N., D. Marraud et H. Sahbi. 2012a, « Camera pose estimation using visual servoing for aerial video change detection », dans *Proceedings of the IEEE International Geoscience And Remote Sensing Symposium*, IEEE Computer Society, p. 3459–3462
- 5 Bourdis, N., D. Marraud et H. Sahbi. 2012b, « Spatio temporal interaction for aerial video change detection », dans *Proceedings of the IEEE International Geoscience And Remote Sensing Symposium*, IEEE Computer Society, p. 2253–2256
- 6 Démonstrateur, basé sur Qt, OpenGL et OpenCV

En cours

- 7 Bourdis, N. et D. Marraud. 2013, « Procédé d'analyse de régions géographiques et de détection de zones d'intérêt », Brevet
- 8 Bourdis, N., H. Sahbi et D. Marraud. 2013b, « Reliable detection of significant changes in aerial videos with arbitrary trajectories », *IEEE Transactions on Geoscience and Remote Sensing*
- 9 Bourdis, N., D. Marraud et H. Sahbi. 2013a, « Interactive consolidation of aerial video change detection using relevance feedback »,