

Séparation de sources et détection d'anomalies en imagerie Hyperspectrale

Alexis Huck, Magellium

Mireille Guillaume, Institut Fresnel

*Colloque d'inauguration de la Société Française de
Téledétection Hyperspectrale
7-8 avril 2011, Paris*



Problématique

Modèle des données

Résultats expérimentaux

Conclusion

Table of contents

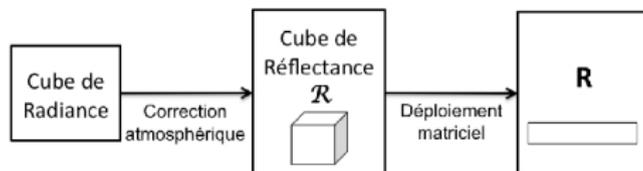
Problématique

Modèle des données

Résultats expérimentaux

Conclusion

Séparation de sources et détection

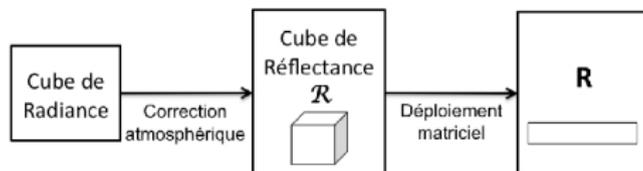


Séparation de sources

$$\mathbf{R} = \mathbf{A}\mathbf{S}$$

- ▶ Démixage
- ▶ Analyse en composantes indépendantes
- ▶ Poursuite de projections

Séparation de sources et détection



Séparation de sources

$$\mathbf{R} = \mathbf{AS}$$

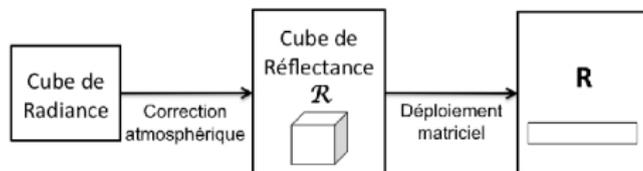
- ▶ Démixage
- ▶ Analyse en composantes indépendantes
- ▶ Poursuite de projections

Détection d'objets

$$\mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

- ▶ LRT ou GLRT
- ▶ Supervisée, non supervisée
- ▶ Taux de fausse alarme constant ?

Séparation de sources et détection



Séparation de sources

$$\mathbf{R} = \mathbf{AS}$$

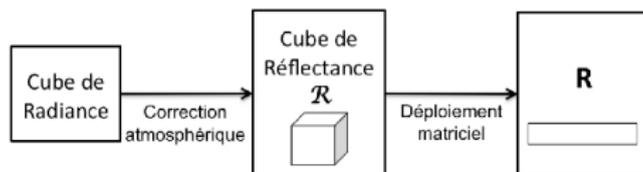
- ▶ Démixage
- ▶ Analyse en composantes indépendantes
- ▶ Poursuite de projections

Détection d'objets

$$\mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

- ▶ LRT ou GLRT
- ▶ Supervisée, non supervisée
- ▶ Taux de fausse alarme constant ?

Séparation de sources et détection



Séparation de sources

$$\mathbf{R} = \mathbf{AS}$$

- ▶ Démixage
- ▶ Analyse en composantes indépendantes
- ▶ **NON!** anomalies ? →
- ▶ Poursuite de projections →

Détection d'objets

$$\mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

- ▶ LRT ou GLRT
- ▶ Supervisée, non supervisée
- ▶ Taux de fausse alarme constant ?

Anomalies en imagerie hyperspectrale

Définition

- ▶ 'Objets' rares et de petite taille (distribution parcimonieuse des abondances)
- ▶ Spectralement distincts du 'fond'

Détection des anomalies

→ Test d'hypothèses (GLRT):

$$\Lambda(\mathbf{r}) \equiv \frac{p(\mathbf{r}; \hat{\theta}_1 | H_1)}{p(\mathbf{r}; \hat{\theta}_0 | H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$$

→ Algorithme 'RX' et variantes (Reed and Yu, 1990)

$$(\mathbf{r} - \mu)' \hat{\Gamma}^{-1} (\mathbf{r} - \mu)$$

→ TFAC, non discriminant

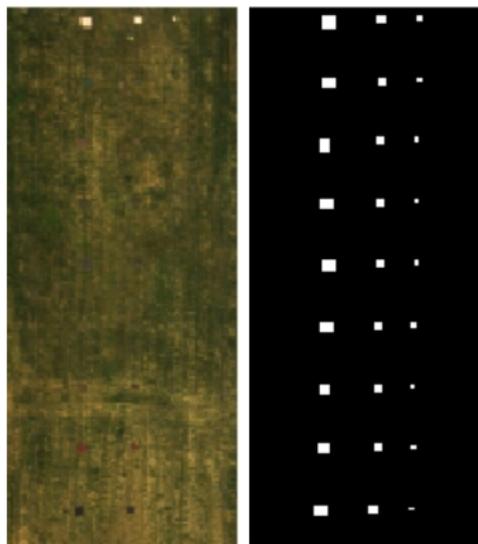


Poursuite de projections

- ▶ Recherche de directions optimisant un critère
- ▶ Critère sensible aux anomalies → directions 'anormales'¹
- ▶ Permet de discriminer entre anomalies
- ▶ Pas de détection TFAC (carte de détection ?)

¹Q. Du and I. Kopriva, 'Automated target detection and discrimination using constrained kurtosis maximization', IEEE GRSL, 2008

Problématique



Scène (HYDICE) et Vérité terrain

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{cible absente} \rightarrow r \sim \mathcal{N}(m, \Gamma)$$

$$H_1: \text{cible présente} \rightarrow r \sim \mathcal{N}(t^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{cible absente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

$$H_1: \text{cible présente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{t}^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;
- ▶ H_3 : sensibilité des anomalies au kurtosis;
- ▶ H_4 : orthogonalité des spectres des différentes classes de cibles (dans l'espace blanchi);

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{cible absente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

$$H_1: \text{cible présente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{t}^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;
- ▶ H_3 : sensibilité des anomalies au kurtosis;
- ▶ H_4 : orthogonalité des spectres des différentes classes de cibles (dans l'espace blanchi);

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{ cible absente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

$$H_1: \text{ cible présente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{t}^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;
- ▶ H_3 : sensibilité des anomalies au kurtosis;
- ▶ H_4 : orthogonalité des spectres des différentes classes de cibles (dans l'espace blanchi);

Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{ cible absente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

$$H_1: \text{ cible présente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{t}^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;
- ▶ H_3 : sensibilité des anomalies au kurtosis;
- ▶ H_4 : orthogonalité des spectres des différentes classes de cibles (dans l'espace blanchi);



Modèle et hypothèses

Objectifs

- ▶ Détection à Taux de Fausses Alarmes Constant (TFAC);
- ▶ Discrimination des anomalies en classes homogènes;

Modèle et hypothèses

$$H_0: \text{ cible absente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{m}, \Gamma)$$

$$H_1: \text{ cible présente} \quad \rightarrow \quad \mathbf{r} \sim \mathcal{N}(\mathbf{t}^d, \Gamma)$$

- ▶ H_2 : rareté des cibles;
- ▶ H_3 : sensibilité des anomalies au kurtosis;
- ▶ H_4 : orthogonalité des spectres des différentes classes de cibles (dans l'espace blanchi);

Table of contents

Problématique

Modèle des données

Résultats expérimentaux

Conclusion

Modèle usuel

$$\begin{aligned} H_0: \quad \mathbf{r}_\zeta &= \mathbf{b}_\zeta^d \\ H_1: \quad \mathbf{r}_\zeta &= \mathbf{b}_\zeta^d + (\mathbf{t}^d - \mathbf{m}) \end{aligned}$$

$\mathbf{b}_\zeta^d \sim \mathcal{N}(\mathbf{m}, \Gamma)$ vecteur aléatoire 'fond'

Modèle usuel

$$\begin{aligned} H_0: \quad \mathbf{r}_\zeta &= \mathbf{b}_\zeta^d \\ H_1: \quad \mathbf{r}_\zeta &= \mathbf{b}_\zeta^d + (\mathbf{t}^d - \mathbf{m}) \end{aligned}$$

$\mathbf{b}_\zeta^d \sim \mathcal{N}(\mathbf{m}, \Gamma)$ vecteur aléatoire 'fond'

Modèle proposé avant et après blanchiment

$$\mathbf{r}_\zeta = \mathbf{b}_\zeta^d + \beta_\zeta (\mathbf{t}^d - \mathbf{m}) \quad i.e. \quad \mathbf{z}_\zeta = \mathbf{b}_\zeta + \beta_\zeta \mathbf{t}$$

Où β_ζ suit une loi de Bernoulli de paramètre p

$$\beta_\zeta \sim \mathcal{B}(p)$$

p : probabilité de présence d'une anomalie en un pixel.

$p \ll 1$

Projection des données

On cherche une direction de projection 'représentative' de l'anomalie

$$\mathbf{s} = \mathbf{w}^T \mathbf{Z}$$

Variable aléatoire modélisée par

$$s_{\zeta \mathbf{w}} = b_{\zeta \mathbf{w}} + \beta_{\zeta} t_{\mathbf{w}}$$

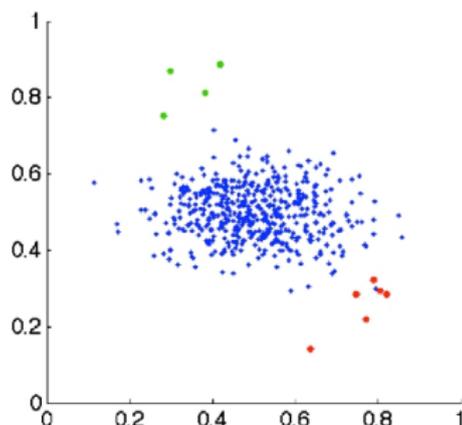
Où $b_{\zeta \mathbf{w}} = \mathbf{w}^T \mathbf{b}_{\zeta} \sim \mathcal{N}(0, 1)$ et $t_{\mathbf{w}} = \mathbf{w}^T \mathbf{t}$

Théorème

\mathbf{w} parallèle à $\mathbf{t} \Rightarrow \mathbf{w}$ maximise (au moins localement) le kurtosis de $s_{\zeta \mathbf{w}}$.

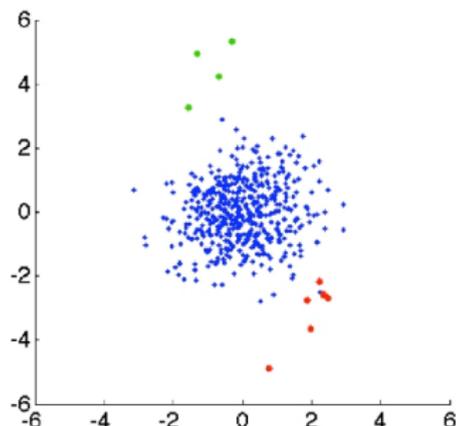
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



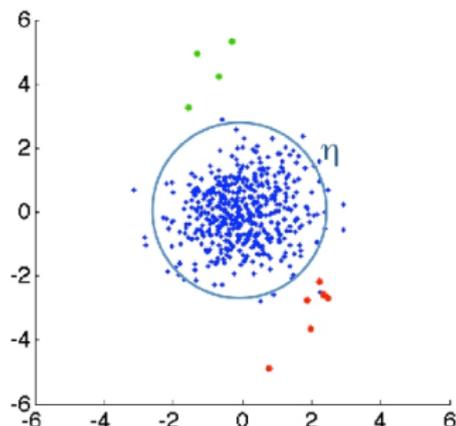
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiment $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de \hat{t}^*
8. Si $\hat{t}^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



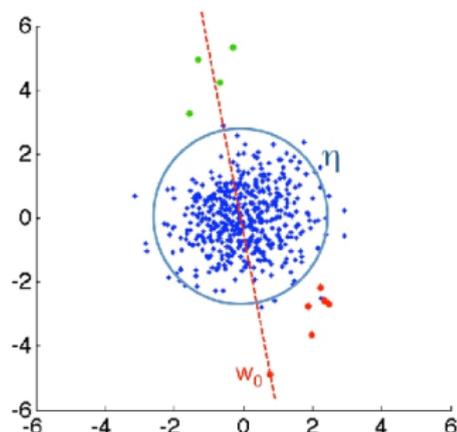
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de \hat{r}^*
8. Si $\hat{r}^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



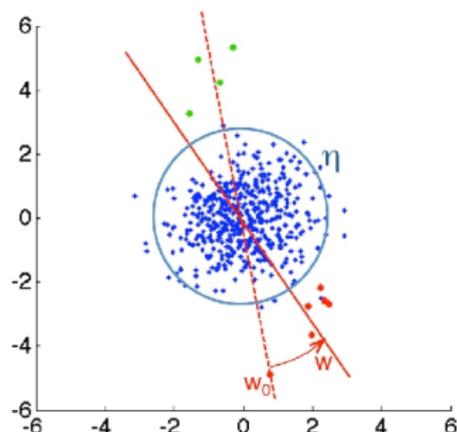
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de \hat{r}^*
8. Si $\hat{r}^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



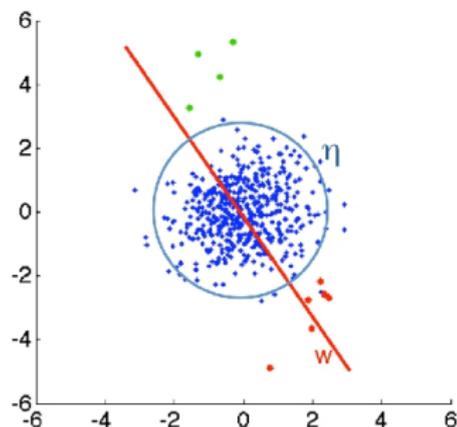
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



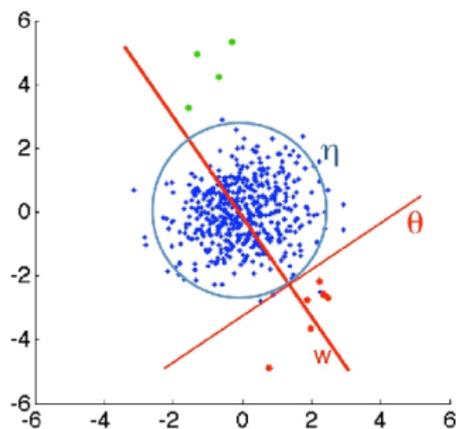
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



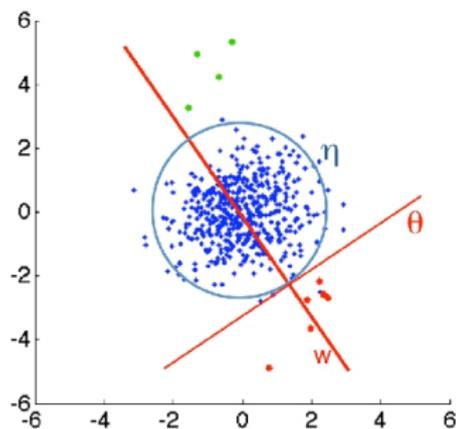
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



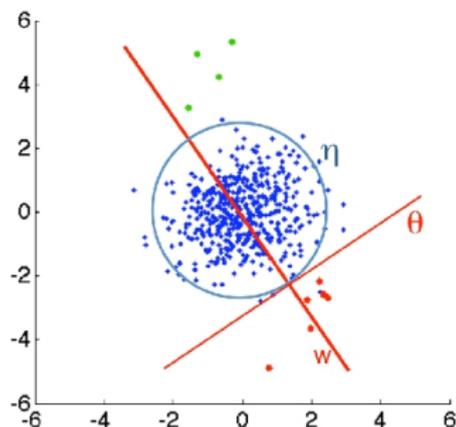
Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



Méthode: *Poursuite de Composantes Anormales* (PCA)

1. $j = 1$ et blanchiement $\rightarrow \mathbf{Z}$
2. MCI (Masque de Conditions Initiales)
3. Projecteur initial: \mathbf{w}_0
4. Maximisation du kurtosis: \mathbf{w}
5. Projection des données sur \mathbf{w}
6. Obtention du masque j
7. $\mathbf{Z} = \mathbf{P}_w^\perp \mathbf{Z}$ puis recherche de i^*
8. Si $i^* \in \text{MCI}$, $j = j + 1$ et retour à 3.



Poursuite de Composantes Anormales

Taux de fausse alarme asymptotiquement constant

$$\blacktriangleright \text{RX} : \Lambda(\mathbf{r}) =_{\infty} \|\mathbf{z}\|_2^2 \underset{H_0}{\overset{H_1}{\geq}} \eta$$

Poursuite de Composantes Anormales

Taux de fausse alarme asymptotiquement constant

▶ RX : $\Lambda(\mathbf{r}) =_{\infty} \|\mathbf{z}\|_2^2 \underset{H_0}{\geq} \underset{H_1}{\leq} \eta$

▶ Taux de fausse alarme $\kappa = \int_{\eta}^{\infty} f_{\chi_L^2}(x) dx$

Poursuite de Composantes Anormales

Taux de fausse alarme asymptotiquement constant

- ▶ RX : $\Lambda(\mathbf{r}) =_{\infty} \|\mathbf{z}\|_2^2 \underset{H_0}{\underset{H_1}{\geq}} \eta$
- ▶ Taux de fausse alarme $\kappa = \int_{\eta}^{\infty} f_{\chi_L^2}(x) dx$
- ▶ Poursuite de la direction la plus anormale autour du pixel le plus anormal $i^* = \operatorname{argmax}_i(\Lambda(\mathbf{z}_i))$

Poursuite de Composantes Anormales

Taux de fausse alarme asymptotiquement constant

- ▶ RX : $\Lambda(\mathbf{r}) =_{\infty} \|\mathbf{z}\|_2^2 \underset{H_0}{\geq} \underset{H_1}{\leq} \eta$
- ▶ Taux de fausse alarme $\kappa = \int_{\eta}^{\infty} f_{\chi_L^2}(x) dx$
- ▶ Poursuite de la direction la plus anormale autour du pixel le plus anormal $i^* = \operatorname{argmax}_i (\Lambda(\mathbf{z}_i))$
- ▶ Projection sur la direction trouvée et test d'hypothèse:
 $\kappa = \int_{\theta}^{\infty} f_{\mathcal{N}(0,1)}(x) dx$

Poursuite de Composantes Anormales

Taux de fausse alarme asymptotiquement constant

- ▶ RX : $\Lambda(\mathbf{r}) =_{\infty} \|\mathbf{z}\|_2^2 \underset{H_0}{\underset{H_1}{\geq}} \eta$
- ▶ Taux de fausse alarme $\kappa = \int_{\eta}^{\infty} f_{\chi_L^2}(x) dx$
- ▶ Poursuite de la direction la plus anormale autour du pixel le plus anormal $i^* = \operatorname{argmax}_i(\Lambda(\mathbf{z}_i))$
- ▶ Projection sur la direction trouvée et test d'hypothèse:
 $\kappa = \int_{\theta}^{\infty} f_{\mathcal{N}(0,1)}(x) dx$

→ j masques de détection

Table of contents

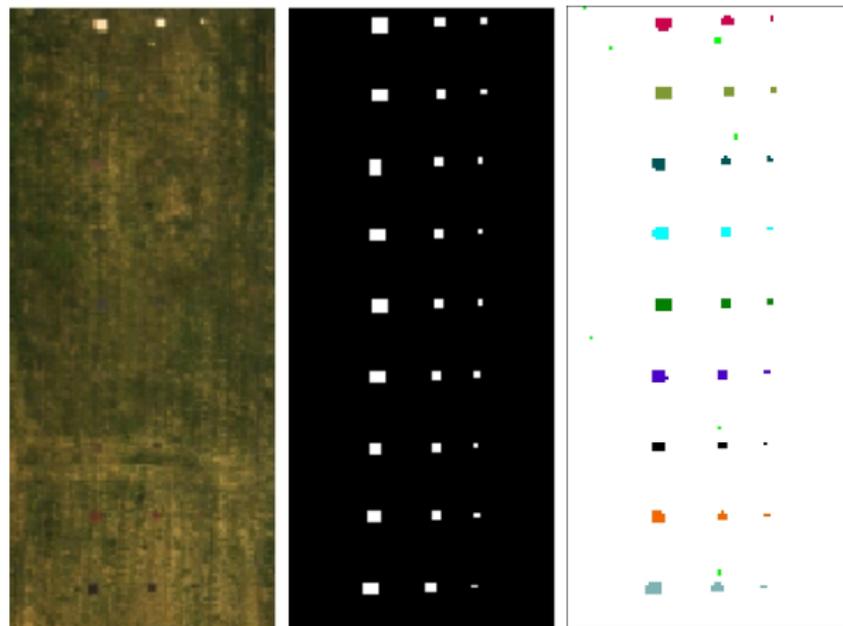
Problématique

Modèle des données

Résultats expérimentaux

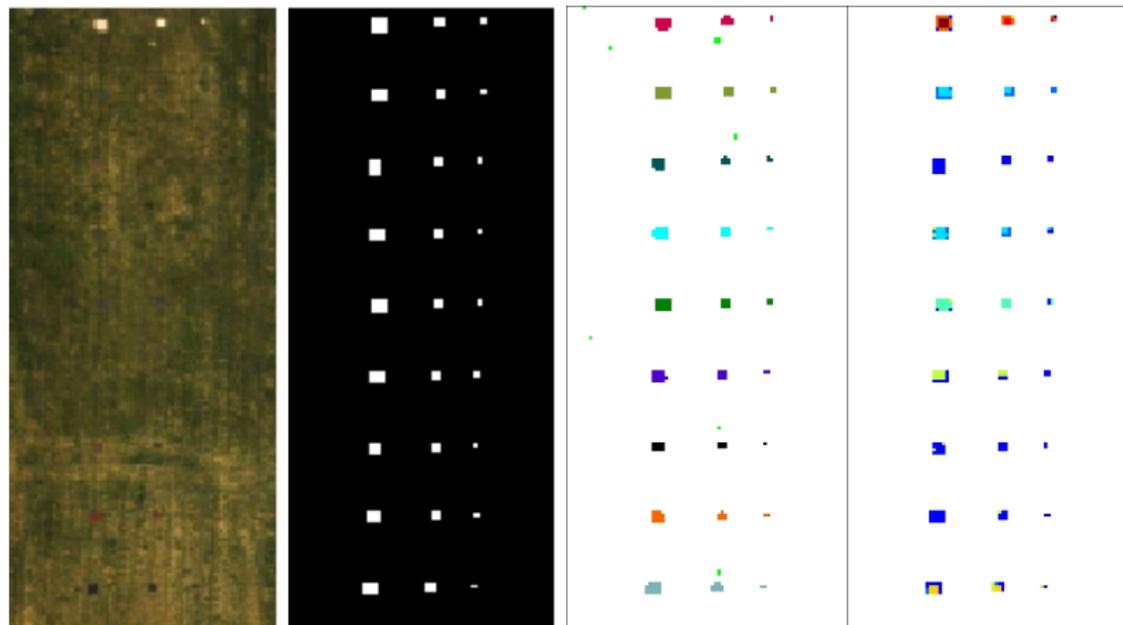
Conclusion

Résultats pour des données HYDICE



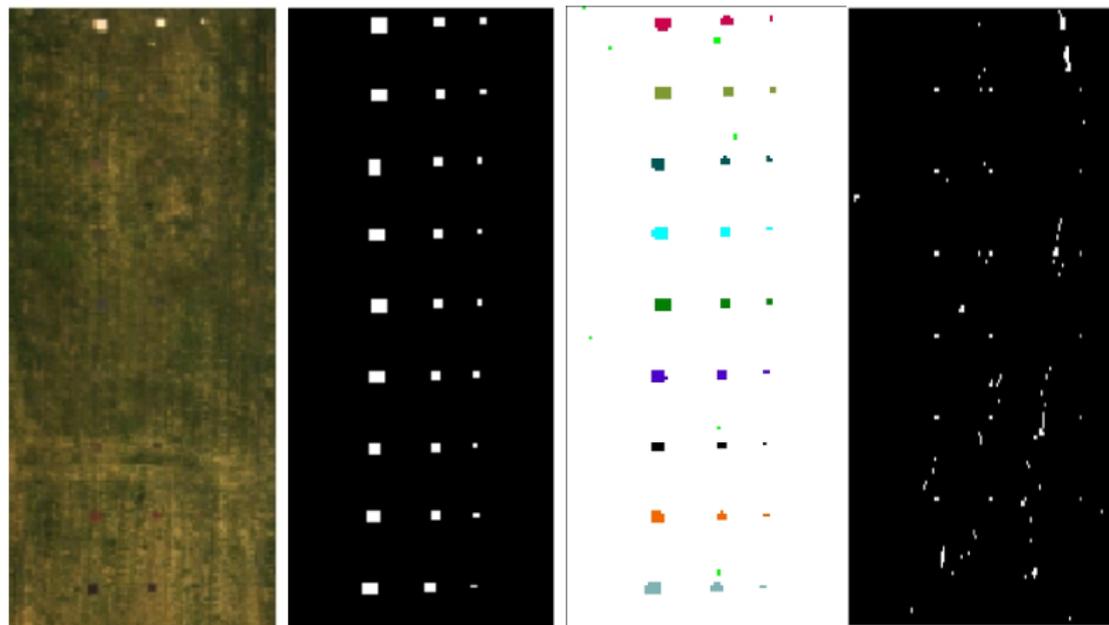
Scène HYDICE / Vérité terrain / P. C. Anormales ($\kappa = 10^{-6}$) / K-means

Résultats pour des données HYDICE



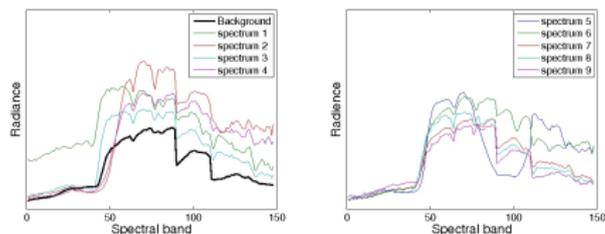
Scène HYDICE / Vérité terrain / P. C. Anormales ($\kappa = 10^{-6}$) / K-means

Résultats pour des données HYDICE

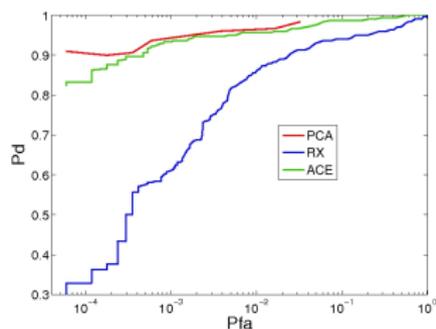


Scène HYDICE / Vérité terrain / P. C. Anormales ($\kappa = 10^{-6}$) / RX

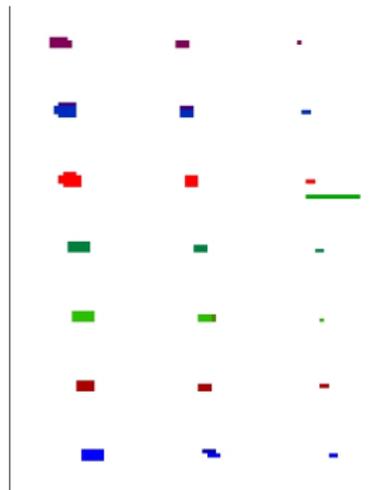
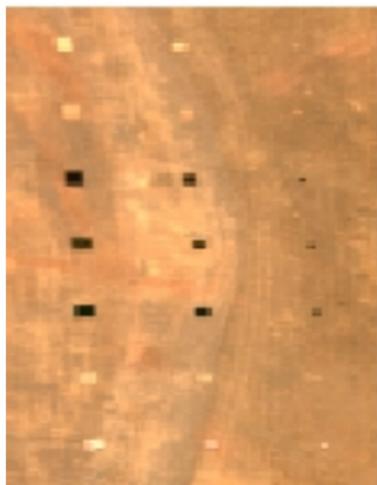
Courbes COR pour les données HYDICE



Spectres du fond et des panneaux

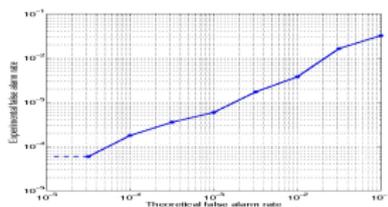


Courbes COR des algorithmes PCA (non supervisé), ACE (supervisé) et RX (non supervisé)



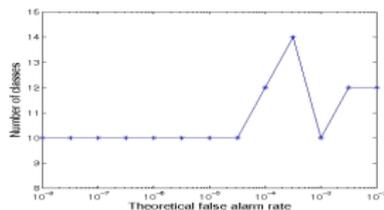
Desert Radiance et carte de détection obtenue avec la Poursuite de Composantes Anormales

Taux de fausse alarme



Taux de fausse alarme théorique en fonction du taux de fausse alarme expérimental

→ taux de fausse alarme asymptotiquement constant



Nombre de composantes en fonction du taux de fausse alarme théorique

Cas de fonds inhomogènes

Image issue de données Cuprite et carte de détection issue du RX

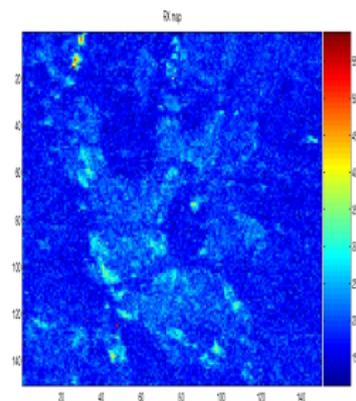
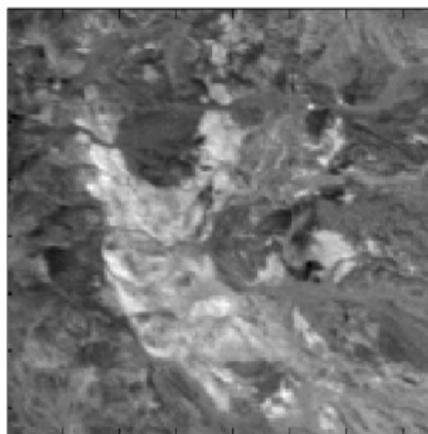
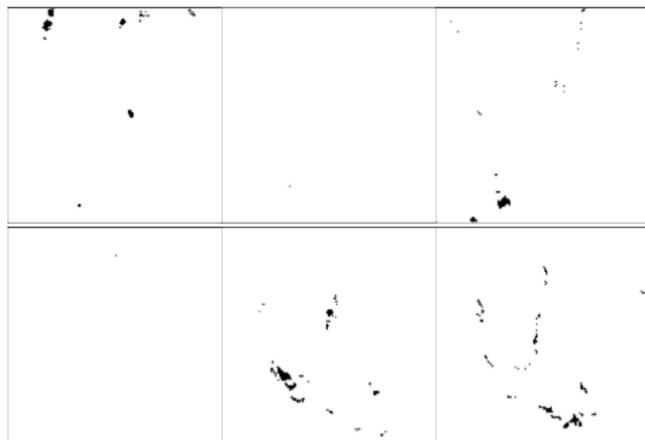
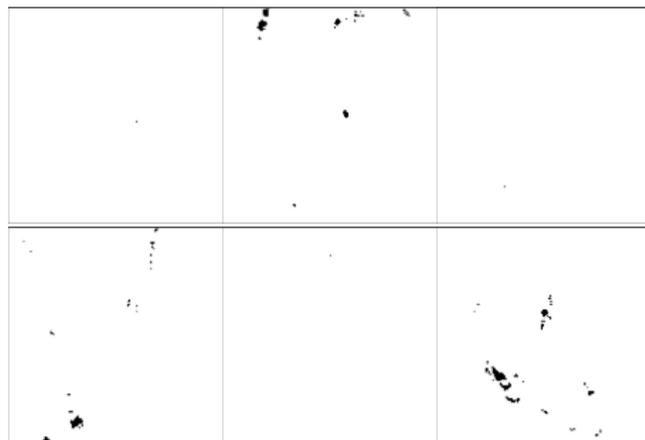


image moyenne de la scene et carte RX

Données Cuprite $Pfa = 10^{-6}$



Six premiers masques d'anomalies

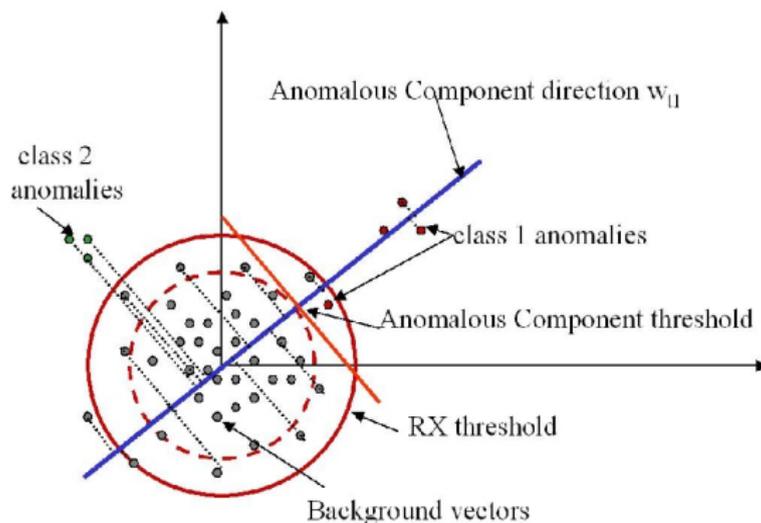


Une anomalie supplémentaire

Interprétation des résultats

Meilleure détection à taux de fausse alarme fixé

→ dimensionalité



Représentation schématique de la poursuite de composantes

Table of contents

Problématique

Modèle des données

Résultats expérimentaux

Conclusion

Conclusion PCA². SYNERGIE I

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
 - un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA² SYNERGIE I

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
 - un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA² SYNERGIE I

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection a taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA² SYNERGIE I

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA². SYNERGIE !

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA²: SYNERGIE !

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA²: SYNERGIE !

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;

Conclusion PCA²: SYNERGIE !

- ▶ Détection basée sur des Tests d'Hypothèses (TH) probabilistes et sur la Poursuite de Projections (PP);
- ▶ Exploitation conjointe des statistiques d'ordre 2 et 4;
- ▶ Propriété TFAC asymptotique
- ▶ Pouvoir discriminant:
un masque de détection par famille d'anomalies
- ▶ Meilleure détection à taux de fausse alarme fixé
- ▶ Estimation automatique du nombre de *composantes anormales*;