Chaînes de Markov triplets et Segmentation non supervisée de signaux

Pierre Lanchantin

GET/INT, département CITI

5 décembre 2006

1. Introduction
2. Modélisation de la loi jointe $Z = (X, Y)$
3. Chaînes de Markov M-non stationnaires cachées
4. Chaînes de Markov évidentielles cachées
5. Arbres de Markov cachés flous multispectraux
6. Conclusions et perspectives

Principe de la segmentation

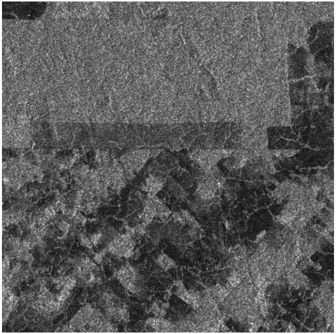


Fig.: Image observée y

Principe de la segmentation

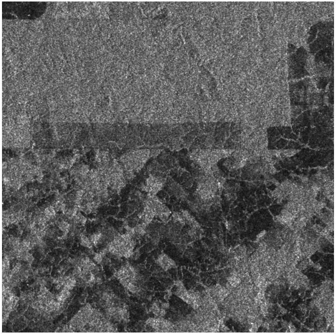


Fig.: Image observée y

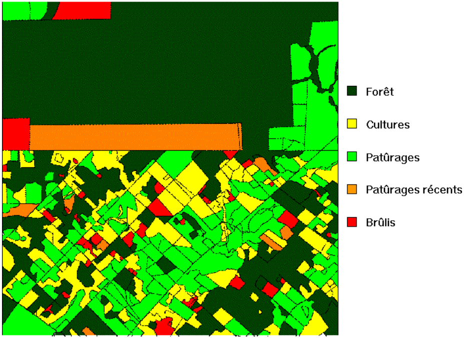


Fig.: Image segmentée x

Principe de la segmentation

Créer une partition x du signal observé y en sous-ensembles appelés *régions* possédant certaines caractéristiques d'homogénéité.

Segmentation statistique bayésienne

Modélisation probabiliste

- $Y = (Y_n)_{n \in \mathcal{N}}$: le p.a observé prenant ses valeurs dans $\mathcal{Y}^N$
- $X = (X_n)_{n \in \mathcal{N}}$: le p.a caché prenant ses valeurs dans $\mathcal{X}^N$
- Liens statistiques modélisés par la loi de $Z = (X, Y)$: $p(z|\phi)$

Segmentation statistique bayésienne

Modélisation probabiliste

- $Y = (Y_n)_{n\in\mathcal{N}}$: le p.a observé prenant ses valeurs dans $\mathcal{Y}^N$
- $X = (X_n)_{n\in\mathcal{N}}$: le p.a caché prenant ses valeurs dans $\mathcal{X}^N$
- Liens statistiques modélisés par la loi de $Z = (X, Y)$: $p(z|\phi)$

- Objectif : Estimer une réalisation $\hat{x}$ à partir de y

Estimateur du mode des marginales a posteriori (MPM)

Minimise le taux de sites mal classés : $\forall n \in \mathcal{N}$,

$$(\hat{x}_n)_{MPM} = \arg\max_{x'_n \in \mathcal{X}} p(x'_n|y, \phi)$$

Segmentation statistique bayésienne

Modélisation probabiliste

- $Y = (Y_n)_{n \in \mathcal{N}}$: le p.a observé prenant ses valeurs dans $\mathcal{Y}^N$
- $X = (X_n)_{n \in \mathcal{N}}$: le p.a caché prenant ses valeurs dans $\mathcal{X}^N$
- Liens statistiques modélisés par la loi de $Z = (X, Y)$: $p(z|\phi)$

- Objectif : Estimer une réalisation $\hat{x}$ à partir de y

Estimateur du mode des marginales a posteriori (MPM)

Minimise le taux de sites mal classés : $\forall n \in \mathcal{N}$,

$$(\hat{x}_n)_{MPM} = \arg\max_{x'_n \in \mathcal{X}} p(x'_n | y, \phi)$$

- Les paramètres ϕ doivent être au préalable estimés (algorithme Espérance Maximisation (EM)) : segmentation non supervisée

Segmentation statistique bayésienne

Modélisation probabiliste

- $Y = (Y_n)_{n \in \mathcal{N}}$: le p.a observé prenant ses valeurs dans $\mathcal{Y}^N$
- $X = (X_n)_{n \in \mathcal{N}}$: le p.a caché prenant ses valeurs dans $\mathcal{X}^N$
- Liens statistiques modélisés par la loi de $Z = (X, Y)$: $p(z|\phi)$

- Objectif : Estimer une réalisation $\hat{x}$ à partir de y

Estimateur du mode des marginales a posteriori (MPM)

Minimise le taux de sites mal classés : $\forall n \in \mathcal{N}$,

$$(\hat{x}_n)_{MPM} = \arg\max_{x'_n \in \mathcal{X}} p(x'_n|y, \phi)$$

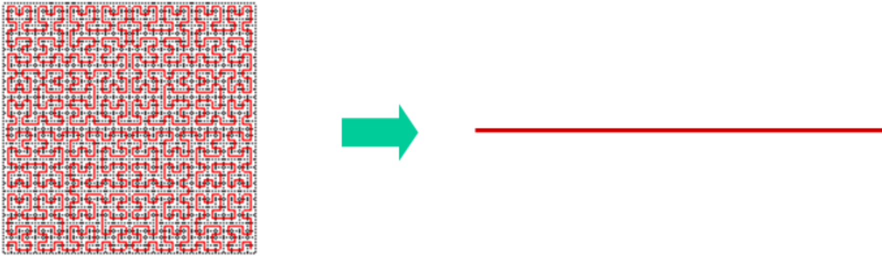
- Les paramètres ϕ doivent être au préalable estimés (algorithme Espérance Maximisation (EM)) : segmentation non supervisée

Choix de la loi jointe $Z = (X, Y)$

- Définir la loi de Z tel que les $p(x_n|y, \phi)$ $\forall n \in \mathcal{N}$ soient calculables
- Le modèle doit être suffisamment riche pour pouvoir modéliser un grand nombre de situations et de comportements

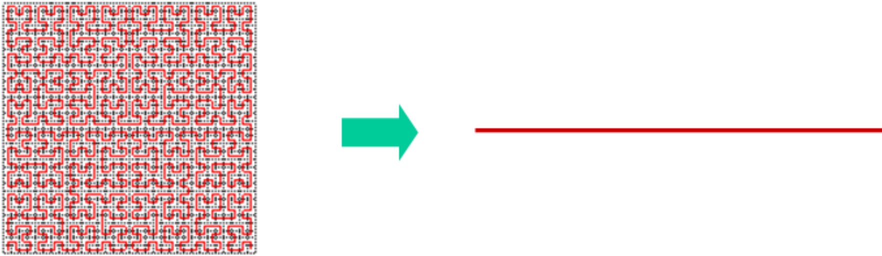
Conversion d'image 2D en processus monodimensionnel

- Conversion 2D -> 1D réalisée à partir d'un parcours de Hilbert-Peano



Conversion d'image 2D en processus monodimensionnel

- Conversion 2D -> 1D réalisée à partir d'un parcours de Hilbert-Peano



- Permet la mise en place d'algorithmes de calcul récursif des probabilités *a posteriori* nécessaires à l'estimation du processus caché.

Calcul des probabilités marginales *a posteriori* dans le cas général

Ecritures factorisées des lois de $Z = (X, Y)$ et de $X|Y$

$$p(z) = p(z_1) \prod_{n=2}^{N} p(z_n|z_{1:n-1}) \text{ et } p(x|y) = p(x_1|y) \prod_{n=2}^{N} p(x_n|x_{1:n-1}, y)$$

Calcul des probabilités marginales *a posteriori* dans le cas général

Ecritures factorisées des lois de $Z = (X, Y)$ et de $X|Y$

$$p(z) = p(z_1) \prod_{n=2}^{N} p(z_n|z_{1:n-1}) \text{ et } p(x|y) = p(x_1|y) \prod_{n=2}^{N} p(x_n|x_{1:n-1}, y)$$

Calcul des lois de transitions *a posteriori*

- Les lois de transitions *a posteriori* peuvent être écrites sous la forme :

$$p(x_n|x_{1:n-1}, y) \propto p(z_n|z_{1:n-1})\beta_n(z_{1:n})$$

Calcul des probabilités marginales *a posteriori* dans le cas général

Ecritures factorisées des lois de $Z = (X, Y)$ et de $X|Y$

$$p(z) = p(z_1) \prod_{n=2}^{N} p(z_n|z_{1:n-1}) \text{ et } p(x|y) = p(x_1|y) \prod_{n=2}^{N} p(x_n|x_{1:n-1}, y)$$

Calcul des lois de transitions *a posteriori*

- Les lois de transitions *a posteriori* peuvent être écrites sous la forme :

$$p(x_n|x_{1:n-1}, y) \propto p(z_n|z_{1:n-1})\beta_n(z_{1:n})$$

- avec $\beta_n(z_{1:n}) = p(y_{n+1:N}|z_{1:n})$ exprimables récursivement $\forall n \in \mathcal{N} \setminus \{1\}$:

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

Calcul des probabilités marginales *a posteriori* dans le cas général

Ecritures factorisées des lois de $Z = (X, Y)$ et de $X|Y$

$$p(z) = p(z_1) \prod_{n=2}^{N} p(z_n|z_{1:n-1}) \text{ et } p(x|y) = p(x_1|y) \prod_{n=2}^{N} p(x_n|x_{1:n-1}, y)$$

Calcul des lois de transitions *a posteriori*

- Les lois de transitions *a posteriori* peuvent être écrites sous la forme :

$$p(x_n|x_{1:n-1}, y) \propto p(z_n|z_{1:n-1})\beta_n(z_{1:n})$$

- avec $\beta_n(z_{1:n}) = p(y_{n+1:N}|z_{1:n})$ exprimables récursivement $\forall n \in \mathcal{N} \setminus \{1\}$:

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

- Cette formulation dans le cas général sera valable pour l'ensemble des modèles présentés par la suite
- Permet la mise en place d'algorithmes de calcul récursif des $p(x_n|y)$ $\forall n \in \mathcal{N}$ en deux passes sur le processus.

Algorithme de calcul des marginales *a posteriori*

◁ Passe arrière : pour $n = N \dots 2$

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n | z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n | z_{1:n-1}) \beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$$p(x_n | x_{1:n-1}, y) \propto \begin{cases} p(z_n | z_{1:n-1}) & \text{si } n = N \\ p(z_n | z_{1:n-1}) \beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

Algorithme de calcul des marginales *a posteriori*

◁ Passe arrière : pour $n = N \dots 2$

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$$p(x_n|x_{1:n-1}, y) \propto \begin{cases} p(z_n|z_{1:n-1}) & \text{si } n = N \\ p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$n = 1$

$$p(x_1|y) \propto p(z_1)\beta_1(z_1)$$

Algorithme de calcul des marginales *a posteriori*

◁ Passe arrière : pour $n = N \dots 2$

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$$p(x_n|x_{1:n-1}, y) \propto \begin{cases} p(z_n|z_{1:n-1}) & \text{si } n = N \\ p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$n = 1$

$$p(x_1|y) \propto p(z_1)\beta_1(z_1)$$

▷ Passe avant : pour $n = 2 \dots N$

$$p(x_{1:n}|y) = p(x_{1:n-1}|y)p(x_n|x_{1:n-1}, y)$$
$$p(x_n|y) = \sum_{x_{1:n-1} \in \mathcal{X}^{n-1}} p(x_{1:n}|y)$$

Algorithme de calcul des marginales *a posteriori*

◁ Passe arrière : pour $n = N \dots 2$

$$\beta_{n-1}(z_{1:n-1}) \propto \begin{cases} \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1}) & \text{si } n = N \\ \sum_{x_n \in \mathcal{X}} p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$$p(x_n|x_{1:n-1}, y) \propto \begin{cases} p(z_n|z_{1:n-1}) & \text{si } n = N \\ p(z_n|z_{1:n-1})\beta_n(z_{1:n}) & \text{sinon} \end{cases}$$

$n = 1$

$$p(x_1|y) \propto p(z_1)\beta_1(z_1)$$

▷ Passe avant : pour $n = 2 \dots N$

$$p(x_{1:n}|y) = p(x_{1:n-1}|y)p(x_n|x_{1:n-1}, y)$$
$$p(x_n|y) = \sum_{x_{1:n-1} \in \mathcal{X}^{n-1}} p(x_{1:n}|y)$$

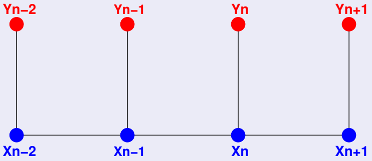
Hypothèses d'indépendance conditionnelle

- $\{p(z_n|z_{1:n-1})\}_{n=2:N}$ non calculables numériquement pour de grandes valeurs de N
- Introduction d'hypothèses d'indépendance conditionnelle.

Chaîne de Markov Cachée à Bruit Indépendant (CMC-BI)

Hypothèses d'indépendance conditionnelle

$$\begin{cases} X_{n+1} \perp\!\!\!\perp X_{1:n-1} | X_n \\ Y_n \perp\!\!\!\perp Z_{\mathcal{N}\setminus n} | X_n \end{cases} \Rightarrow p(z_n | z_{1:n-1}) = p(x_n | x_{n-1}) p(y_n | x_n)$$

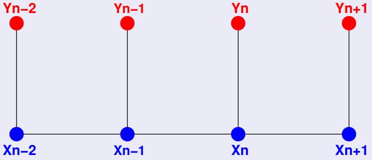


- X est une chaîne de Markov d'ordre 1 (CM).

Chaîne de Markov Cachée à Bruit Indépendant (CMC-BI)

Hypothèses d'indépendance conditionnelle

$$\begin{cases} X_{n+1} \perp\!\!\!\perp X_{1:n-1} | X_n \\ Y_n \perp\!\!\!\perp Z_{\mathcal{N}\setminus n} | X_n \end{cases} \Rightarrow p(z_n|z_{1:n-1}) = p(x_n|x_{n-1})p(y_n|x_n)$$

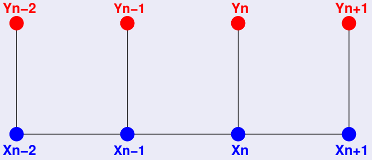


- X est une chaîne de Markov d'ordre 1 (CM).
- Z est une CM

Chaîne de Markov Cachée à Bruit Indépendant (CMC-BI)

Hypothèses d'indépendance conditionnelle

$$\begin{cases} X_{n+1} \perp\!\!\!\perp X_{1:n-1} | X_n \\ Y_n \perp\!\!\!\perp Z_{\mathcal{N}\setminus n} | X_n \end{cases} \Rightarrow p(z_n|z_{1:n-1}) = p(x_n|x_{n-1})p(y_n|x_n)$$

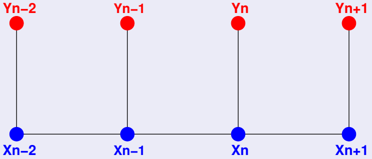


- X est une chaîne de Markov d'ordre 1 (CM).
- Z est une CM
- $X|Y$ est une CM $\rightarrow$ calcul numérique possible des marginales *a posteriori*

Chaîne de Markov Cachée à Bruit Indépendant (CMC-BI)

Hypothèses d'indépendance conditionnelle

$$\begin{cases} X_{n+1} \perp\!\!\!\perp X_{1:n-1} | X_n \\ Y_n \perp\!\!\!\perp Z_{\mathcal{N}\setminus n} | X_n \end{cases} \Rightarrow p(z_n|z_{1:n-1}) = p(x_n|x_{n-1})p(y_n|x_n)$$

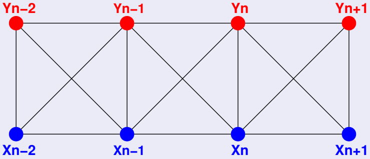


- X est une chaîne de Markov d'ordre 1 (CM).
- Z est une CM
- $X|Y$ est une CM $\rightarrow$ calcul numérique possible des marginales *a posteriori*
- Modèle largement utilisé : parcimonie, grande robustesse.

Chaîne de Markov Couple (CMCo)

Hypothèse d'indépendance conditionnelle

$Z_{n+1} \perp\!\!\!\perp Z_{1:n-1} | Z_n \Rightarrow p(z_n|z_{1:n-1}) = p(z_n|z_{n-1})$

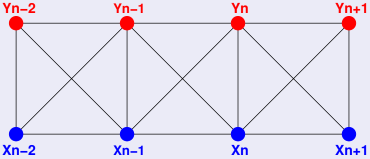


- On suppose directement la markovianité de Z

Chaîne de Markov Couple (CMCo)

Hypothèse d'indépendance conditionnelle

$$Z_{n+1} \perp\!\!\!\perp Z_{1:n-1} | Z_n \Rightarrow p(z_n | z_{1:n-1}) = p(z_n | z_{n-1})$$

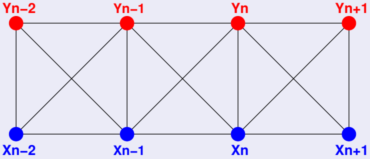


- On suppose directement la markovianité de Z
- $X|Y$ et $Y|X$ sont des CM

Chaîne de Markov Couple (CMCo)

Hypothèse d'indépendance conditionnelle

$$Z_{n+1} \perp\!\!\!\perp Z_{1:n-1} | Z_n \Rightarrow p(z_n | z_{1:n-1}) = p(z_n | z_{n-1})$$

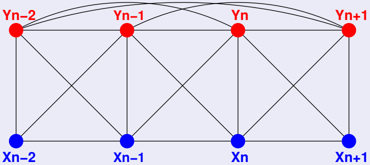


- On suppose directement la markovianité de Z
- $X|Y$ et $Y|X$ sont des CM
- X n'est plus nécessairement une CM : modèle plus général que CMC-BI : on connaît une condition nécessaire et suffisante mettant en lumière l'intérêt des CMCouples par rapport aux CMC-BI

Chaîne Couple Partiellement de Markov (CCPM)

Hypothèse d'indépendance conditionnelle

$$Z_{n+1} \perp\!\!\!\perp X_{1:n-1} | (Z_n, Y_{1:n-1}) \Rightarrow p(z_n | z_{1:n-1}) = p(z_n | z_{n-1}, y_{1:n-2})$$

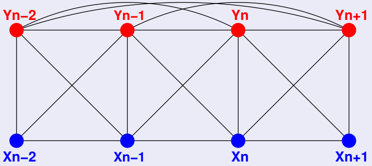


- On obtient directement la markovianité de $X|Y$

Chaîne Couple Partiellement de Markov (CCPM)

Hypothèse d'indépendance conditionnelle

$$Z_{n+1} \perp\!\!\!\perp X_{1:n-1} | (Z_n, Y_{1:n-1}) \Rightarrow p(z_n | z_{1:n-1}) = p(z_n | z_{n-1}, y_{1:n-2})$$



- On obtient directement la markovianité de $X|Y$
- $Y|X$ n'est pas nécessairement une CM : prise en compte de bruits à corrélation longue tels que :

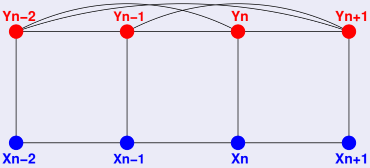
$$\lim_{\tau \to \infty} \tau^{\alpha} \rho(\tau) = c_{\rho}$$

où $\alpha \in]0, 1[$, $c_{\rho} > 0$ est une constante, et $\rho(\tau)$ est une fonction d'autocorrélation.

Chaîne de Markov Cachée à bruit à corrélation longue (CMC-BCL)

Cas particulier de CCPM gaussienne avec

$$\begin{cases} X_{n+1} \perp\!\!\!\perp X_{1:n-1} | X_n \\ Y_n \perp\!\!\!\perp X_{\mathcal{N}\setminus n} | X_n,\, Y_{1:n-1} \end{cases} \Rightarrow p(z_n|z_{1:n-1}) = p(x_n|x_{n-1})p(y_{1:n}|x_n)$$



Restauration supervisée d'un processus à bruit à corrélation longue

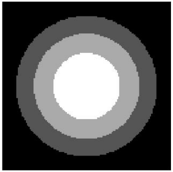


Fig.: Initiale

Restauration supervisée d'un processus à bruit à corrélation longue

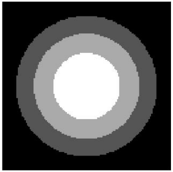


Fig.: Initiale

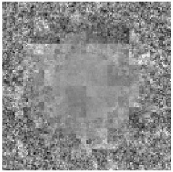


Fig.: Bruitée

- Processus gaussiens stationnaires centrées, variances=1
- Les autocorrélations sont les seuls paramètres discriminants

Restauration supervisée d'un processus à bruit à corrélation longue

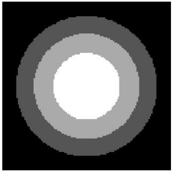


Fig.: Initiale

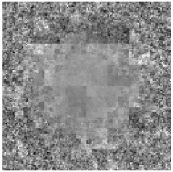


Fig.: Bruitée

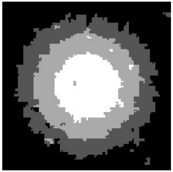


Fig.: Restaurée

- Processus gaussiens stationnaires centrées, variances=1
- Les autocorrélations sont les seuls paramètres discriminants

Autocorrélations

$\rho(\tau) = |\tau + 1|^{-\alpha}$ avec $\tau = |j - i|$ avec $\alpha_{\omega_1} = 0.99$,$\alpha_{\omega_2} = 0.3$,$\alpha_{\omega_3} = 0.05$, $\alpha_{\omega_4} = 0.01$

- MPM : taux d'erreur 6.9%

Segmentation non supervisée d'un processus à bruit à corrélation longue

Fig.: Initiale

- Image à 2 classes
- On souhaite étudier la robustesse des modèles CMC-BI et CMC-BCL

Segmentation non supervisée d'un processus à bruit à corrélation longue

Fig.: Initiale

- Image à 2 classes
- On souhaite étudier la robustesse des modèles CMC-BI et CMC-BCL
- On considère 2 types de bruits :

Segmentation non supervisée d'un processus à bruit à corrélation longue

Fig.: Initiale

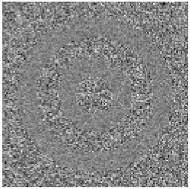


Fig.: BI

- Image à 2 classes
- On souhaite étudier la robustesse des modèles CMC-BI et CMC-BCL
- On considère 2 types de bruits :
  - Bruit indépendant (BI) : variances (1,4) , moyennes (0,0)

Segmentation non supervisée d'un processus à bruit à corrélation longue

Fig.: Initiale

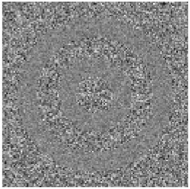


Fig.: BI

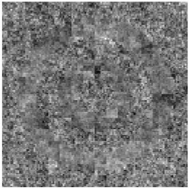


Fig.: BCL

- Image à 2 classes
- On souhaite étudier la robustesse des modèles CMC-BI et CMC-BCL
- On considère 2 types de bruits :
  - Bruit indépendant (BI) : variances (1,4) , moyennes (0,0)
  - Bruit à corrélation longue (BCL) : variances=1 et $\alpha_{\omega_1} = 0.99, \alpha_{\omega_2} = 0.2$

Segmentation non supervisée d'un processus à bruit à corrélation longue

Fig.: Initiale

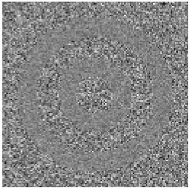


Fig.: BI

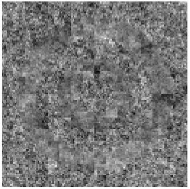


Fig.: BCL

- Image à 2 classes
- On souhaite étudier la robustesse des modèles CMC-BI et CMC-BCL
- On considère 2 types de bruits :
  - Bruit indépendant (BI) : variances (1,4) , moyennes (0,0)
  - Bruit à corrélation longue (BCL) : variances=1 et $\alpha_{\omega_1} = 0.99, \alpha_{\omega_2} = 0.2$
- Paramètres estimés par EM initialisé par K-moyennes en considérant le cas CMC-BI et le cas CMC-BCL

Résultats de segmentations

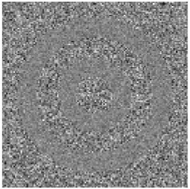


Fig.: Image bruitée par BI

Résultats de segmentations

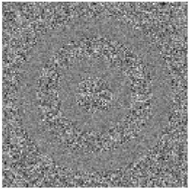


Fig.: Image bruitée par BI

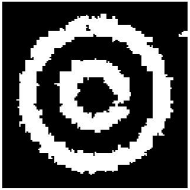


Fig.: CMC-BI (5.3%)

Résultats de segmentations

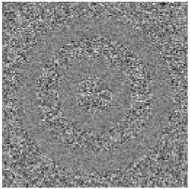


Fig.: Image bruitée par BI

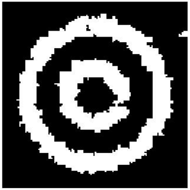


Fig.: CMC-BI (5.3%)

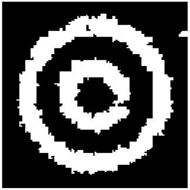


Fig.: CMC-BCL(5.3%)

Résultats de segmentations

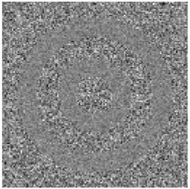


Fig.: Image bruitée par BI

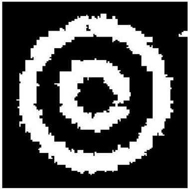


Fig.: CMC-BI (5.3%)

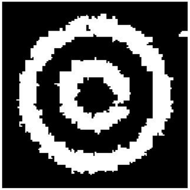


Fig.: CMC-BCL(5.3%)

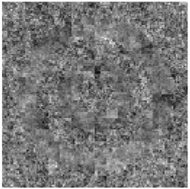


Fig.: Image bruitée par BCL

Résultats de segmentations

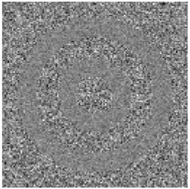


Fig.: Image bruitée par BI

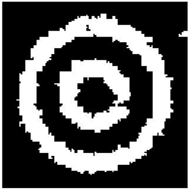


Fig.: CMC-BI (5.3%)

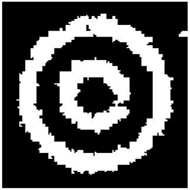


Fig.: CMC-BCL(5.3%)

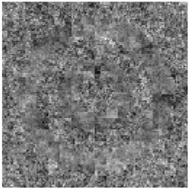


Fig.: Image bruitée par BCL

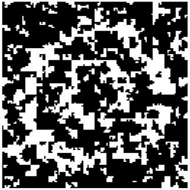


Fig.: CMC-BI(27.6%)

Résultats de segmentations

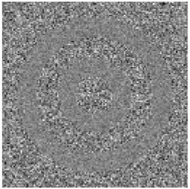


Fig.: Image bruitée par BI

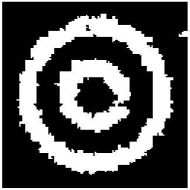


Fig.: CMC-BI (5.3%)

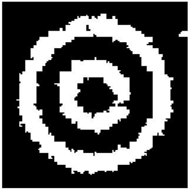


Fig.: CMC-BCL(5.3%)

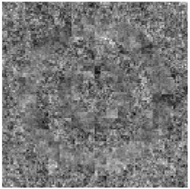


Fig.: Image bruitée par BCL

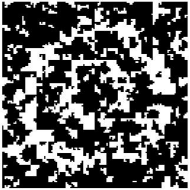


Fig.: CMC-BI(27.6%)

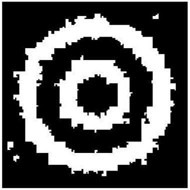


Fig.: CMC-BCL(6.5%)

Principes des modèles triplets partiellement de Markov

Principe des modèles triplets partiellement de Markov

Ajouter un processus auxiliaire caché U prenant ses valeurs dans $\mathcal{U}^N$ discret tel que $V|Y$ soit une CM (avec $V = (X, U)$).

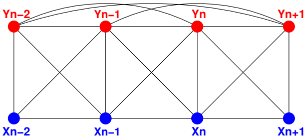


Fig.: CCPM

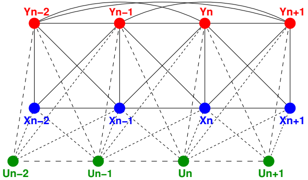


Fig.: CTPM

Principes des modèles triplets partiellement de Markov

Principe des modèles triplets partiellement de Markov

Ajouter un processus auxiliaire caché U prenant ses valeurs dans $\mathcal{U}^N$ discret tel que $V|Y$ soit une CM (avec $V = (X, U)$).

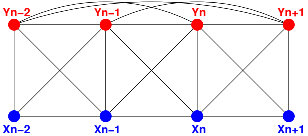


Fig.: CCPM

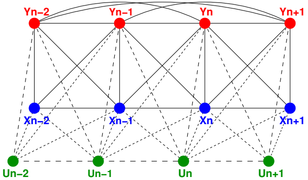


Fig.: CTPM

- $V|y$ étant une CM, est il est possible de calculer les marginales $p(v_n|y)$ et d'en déduire

$$p(x_n|y) = \sum_{u_n \in \mathcal{U}} p(v_n|y) = \sum_{u_n \in \mathcal{U}} p(x_n, u_n|y)$$

Principes des modèles triplets partiellement de Markov

Principe des modèles triplets partiellement de Markov

Ajouter un processus auxiliaire caché U prenant ses valeurs dans $\mathcal{U}^N$ discret tel que $V|Y$ soit une CM (avec $V = (X, U)$).

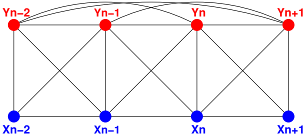


Fig.: CCPM

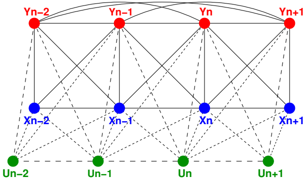


Fig.: CTPM

- $V|y$ étant une CM, est il est possible de calculer les marginales $p(v_n|y)$ et d'en déduire

$$p(x_n|y) = \sum_{u_n \in \mathcal{U}} p(v_n|y) = \sum_{u_n \in \mathcal{U}} p(x_n, u_n|y)$$

- On peut dès lors estimer le processus caché par MPM

Généralisations des modèles présentés

- Chacun des modèles présentés peut être généralisé

Généralisations des modèles présentés

- Chacun des modèles présentés peut être généralisé

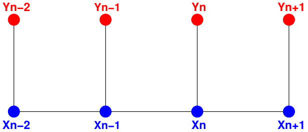


Fig.: CMC-BI

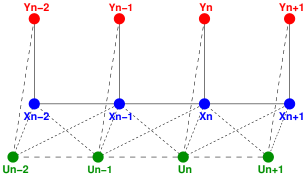


Fig.: CMCouple-BI

Généralisations des modèles présentés

- Chacun des modèles présentés peut être généralisé

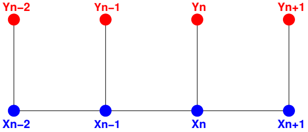


Fig.: CMC-BI

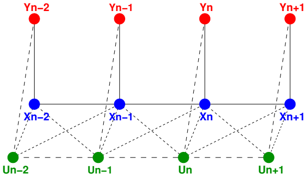


Fig.: CMCouple-BI

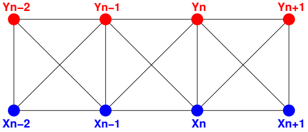


Fig.: CMCouple

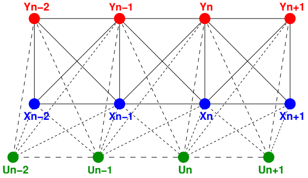


Fig.: CMT

Segmentation de processus non stationnaires

Problématique

- Lors de l'estimation des paramètres ϕ du modèle, $Z = (X, Y)$ est supposé stationnaire.
- Si Z n'est pas stationnaire, mauvaise estimation des paramètres, qui peut se traduire par une mauvaise segmentation.

Segmentation de processus non stationnaires

Problématique

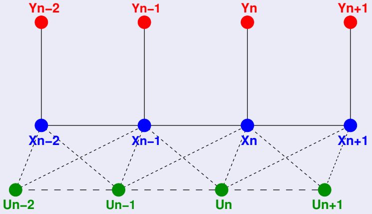
- Lors de l'estimation des paramètres ϕ du modèle, $Z = (X, Y)$ est supposé stationnaire.
- Si Z n'est pas stationnaire, mauvaise estimation des paramètres, qui peut se traduire par une mauvaise segmentation.

Approche proposée

- La non stationnarité de Z est modélisé par un processus auxiliaire U gouvernant les changement de régime du processus.
- On suppose que $T = (X, U, Y)$ est une CMT stationnaire
- $Z|U$ n'est pas stationnaire

Expérimentations

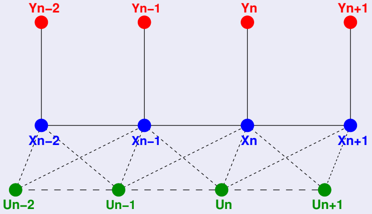
Chaîne de Markov M-Non Stationnaire Cachée BI



- CMT particulière où le processus $V = (X, U)$ est une CM stationnaire et la loi $p(y|v)$ de Y conditionnelle à V vérifie : $p(y|v) = p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$.

Expérimentations

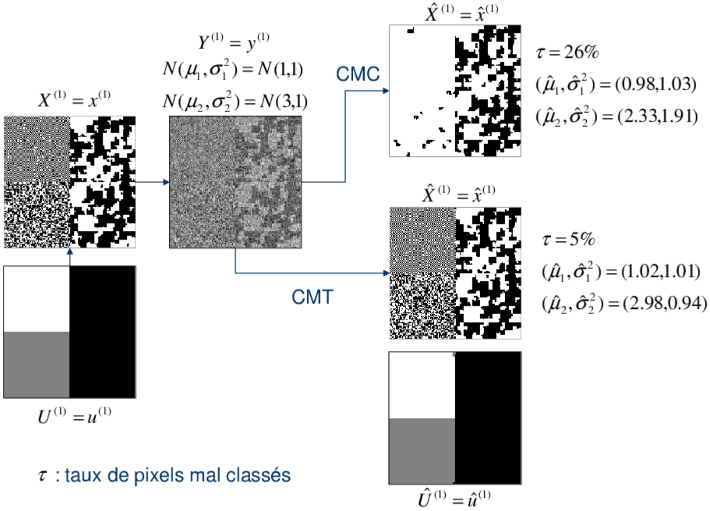
Chaîne de Markov M-Non Stationnaire Cachée BI



- CMT particulière où le processus $V = (X, U)$ est une CM stationnaire et la loi $p(y|v)$ de Y conditionnelle à V vérifie : $p(y|v) = p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$.

- Estimation des paramètres effectuée par Stochastique EM initialisé par un algorithme de K-moyennes
- Le modèle CM M-NSC-BI sera dénoté CMT dans les expérimentations

Segmentation de processus synthétiques



Processus caché réel

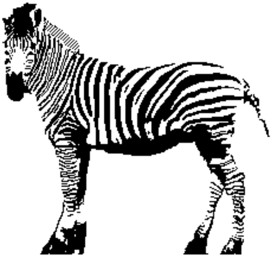


Fig.: Image initiale cachée

- Image binaire
- La peau présente des bandes de tailles et d'espacements variés

Processus caché réel

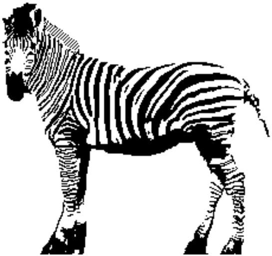


Fig.: Image initiale cachée

Fig.: Image bruitée

- Image binaire
- La peau présente des bandes de tailles et d'espacements variés
- y est échantillonné de la même manière que dans la première expérimentation

Résultats de segmentation

Fig.: CMC-BI(7.2%)

Résultats de segmentation

Fig.: CMC-BI(7.2%)

Fig.: CMT(3.5%)

- Amélioration du résultat de la segmentation

Résultats de segmentation

Fig.: CMC-BI(7.2%)

Fig.: CMT(3.5%)

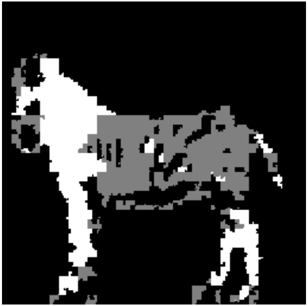


Fig.: $\hat{U} = \hat{u}$

- Amélioration du résultat de la segmentation
- Visualisation des différentes zones sur le corps du zèbre en visualisant l'estimation de $\hat{U} = \hat{u}$

Segmentation de textures de Brodatz

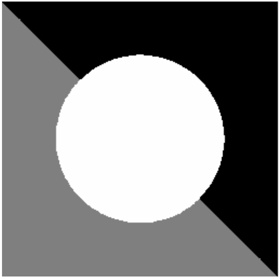


Fig.: Processus auxiliaire initial

Segmentation de textures de Brodatz

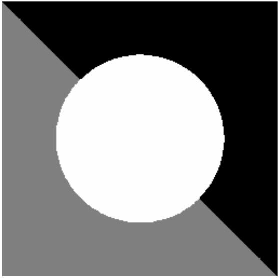


Fig.: Processus auxiliaire initial

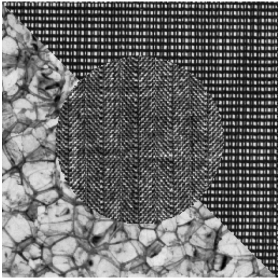


Fig.: Image observée simulée

- y construit par rapport à un processus u en considérant 3 textures différentes de Brodatz

Segmentation de textures de Brodatz

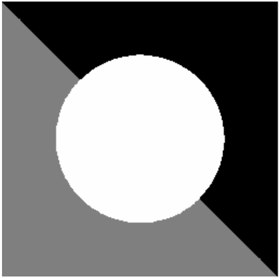


Fig.: Processus auxiliaire initial

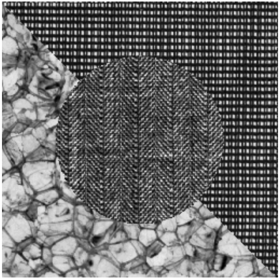


Fig.: Image observée simulée

- y construit par rapport à un processus u en considérant 3 textures différentes de Brodatz
- CMT avec 3 classes auxiliaires et 4 classes réelles

Résultats de segmentation

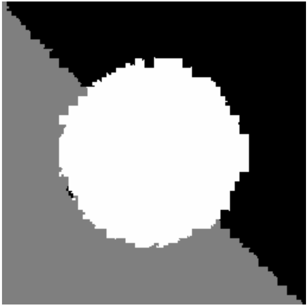


Fig.: processus auxiliaire estimé, $\hat{U} = \hat{u}$

- $\hat{U} = \hat{u}$ est bien estimé avec $\tau = 2.5\%$
- Les CMT peuvent être utiles dans le contexte de la segmentation non supervisée de textures

Résultats de segmentation

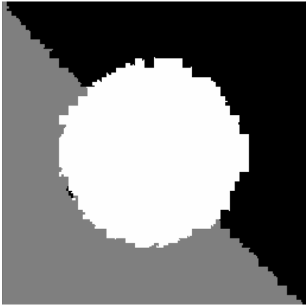


Fig.: processus auxiliaire estimé, $\hat{U} = \hat{u}$

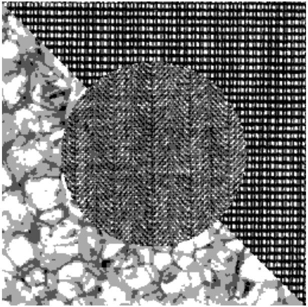


Fig.: processus caché estimé $\hat{X}$

- $\hat{U} = \hat{u}$ est bien estimé avec $\tau = 2.5\%$
- Les CMT peuvent être utiles dans le contexte de la segmentation non supervisée de textures

Segmentation d'une image réelle

Segmentation d'une image réelle

Bayesian Information Criteria (BIC)

$$BIC = -2\log L(y) + 2k\log N$$

$L(y)$: vraisemblance
k : nombre de paramètres non nuls
N : nombre d'échantillons

Valeurs du BIC

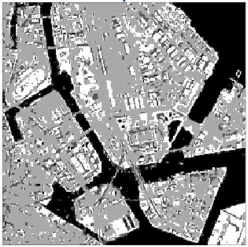
Nombre d'états réels	Nombre d'états auxiliaires	LL	Nombre de paramètres	BIC
2	1	-338750	7	677580
2	2	-337460	13	675060
2	3	-337360	23	**674980**
3	1	-334310	14	668780
3	2	-332070	27	664440
3	3	-331760	55	**664130**
4	1	-333020	21	666280
4	2	-331210	43	**662610**

Segmentation d'image réelle

4 classes réelles

1 état auxiliaire

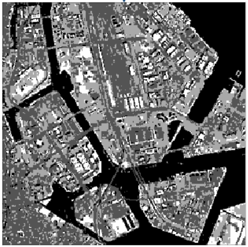
BIC=666280



4 classes réelles

2 états auxiliaires

BIC= 662610



La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer

La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer
- Présentée comme une généralisation de la théorie de Bayes au traitement de l'incertain

La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer
- Présentée comme une généralisation de la théorie de Bayes au traitement de l'incertain

Définition

Une fonction de masse élémentaire m est une fonction de $P(\mathcal{X})$ dans R^+ vérifiant

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in P(\mathcal{X})} m(A) = 1 \end{cases}$$

La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer
- Présentée comme une généralisation de la théorie de Bayes au traitement de l'incertain

Définition

Une fonction de masse élémentaire m est une fonction de $P(\mathcal{X})$ dans R^+ vérifiant

$$\left\{ \begin{array}{l} m\left(\emptyset\right) = 0 \\ \sum_{A \in P(\mathcal{X})} m(A) = 1 \end{array} \right.$$

- Possibilité d'affecter une masse aux hypothèses composites : modélisation plus flexible permettant de représenter l'incomplétude, l'ambiguïté, l'ignorance

La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer
- Présentée comme une généralisation de la théorie de Bayes au traitement de l'incertain

Définition

Une fonction de masse élémentaire m est une fonction de $P(\mathcal{X})$ dans R^+ vérifiant

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in P(\mathcal{X})} m(A) = 1 \end{cases}$$

- Possibilité d'affecter une masse aux hypothèses composites : modélisation plus flexible permettant de représenter l'incomplétude, l'ambiguïté, l'ignorance
- Ignorance totale : $m(\mathcal{X}) = 1$ et $m(A) = 0$, $\forall A \subset \mathcal{X}$, $A \neq \mathcal{X}$

La théorie de L'évidence

- Introduite par Dempster puis reformulée par Shafer
- Présentée comme une généralisation de la théorie de Bayes au traitement de l'incertain

Définition

Une fonction de masse élémentaire m est une fonction de $P(\mathcal{X})$ dans R^+ vérifiant

$$\begin{cases} m(\emptyset) = 0 \\ \sum_{A \in P(\mathcal{X})} m(A) = 1 \end{cases}$$

- Possibilité d'affecter une masse aux hypothèses composites : modélisation plus flexible permettant de représenter l'incomplétude, l'ambiguïté, l'ignorance
- Ignorance totale : $m(\mathcal{X}) = 1$ et $m(A) = 0$, $\forall A \subset \mathcal{X}$, $A \neq \mathcal{X}$
- Jeu de masses bayésien : masses réparties uniquement sur les hypothèses singletons : probabilité

Fusion de Dempster-Shafer (Fusion DS)

Masses $m_1, \ldots, m_l$ fournies par plusieurs sources d'information

Fusion de Dempster-Shafer (Fusion DS)

Masses $m_1, \ldots, m_l$ fournies par plusieurs sources d'information

Définition

$$m(A) = (m_1 \oplus m_2 \oplus \cdots \oplus m_l)(A) \propto \sum_{B_1 \cap \cdots \cap B_l = A} m_1(B_1) m_2(B_2) \ldots m_l(B_l)$$

Fusion de Dempster-Shafer (Fusion DS)

Masses $m_1, \ldots, m_l$ fournies par plusieurs sources d'information

Définition

$$m(A) = (m_1 \oplus m_2 \oplus \cdots \oplus m_l)(A) \propto \sum_{B_1 \cap \cdots \cap B_l = A} m_1(B_1) m_2(B_2) \ldots m_l(B_l)$$

Propriété

Lorsque au moins l'une des masses impliquées dans la combinaison est bayésienne, le résultat de la combinaison est une masse bayésienne.

Fusion de Dempster-Shafer (Fusion DS)

Masses $m_1, \ldots, m_l$ fournies par plusieurs sources d'information

Définition

$$m(A) = (m_1 \oplus m_2 \oplus \cdots \oplus m_l)(A) \propto \sum_{B_1 \cap \cdots \cap B_l = A} m_1(B_1) m_2(B_2) \ldots m_l(B_l)$$

Propriété

Lorsque au moins l'une des masses impliquées dans la combinaison est bayésienne, le résultat de la combinaison est une masse bayésienne.

Calcul de la loi *a posteriori* : cas particulier de Fusion DS.

Soit $m(x) = p(x)$, $x \in \mathcal{X}$, et $q(x) \propto p(y|x)$,$x \in \mathcal{X}$, ($y \in \mathcal{Y}$ est fixé).

$$(m \oplus q)(x) = \frac{m(x)q(x)}{\sum_{x' \in \mathcal{X}} m(x')q(x')} = \frac{p(x)p(y|x)}{\sum_{x' \in \mathcal{X}} p(x')p(y|x')} = p(x|y)$$

Fusion de Dempster-Shafer (Fusion DS)

Masses $m_1, \dots, m_l$ fournies par plusieurs sources d'information

Définition

$$m(A) = (m_1 \oplus m_2 \oplus \cdots \oplus m_l)(A) \propto \sum_{B_1 \cap \cdots \cap B_l = A} m_1(B_1) m_2(B_2) \dots m_l(B_l)$$

Propriété

Lorsque au moins l'une des masses impliquées dans la combinaison est bayésienne, le résultat de la combinaison est une masse bayésienne.

Calcul de la loi *a posteriori* : cas particulier de Fusion DS.

Soit $m(x) = p(x)$, $x \in \mathcal{X}$, et $q(x) \propto p(y|x)$,$x \in \mathcal{X}$, ($y \in \mathcal{Y}$ est fixé).

$$(m \oplus q)(x) = \frac{m(x)q(x)}{\sum_{x' \in \mathcal{X}} m(x')q(x')} = \frac{p(x)p(y|x)}{\sum_{x' \in \mathcal{X}} p(x')p(y|x')} = p(x|y)$$

Généralisation du calcul de la loi a posteriori

Remplacement de l'une des masses par une masse évidentielle $m(u)$, $u \in \mathcal{P}(\mathcal{X})$.

Affaiblissement

Affaiblissement

Probabilité est connue avec imprécision : remplacement de la probabilité non fiable $p(x)$ par la masse $m(u)$ définie par :

$$m(u) = \alpha p(x), \quad m(\mathcal{X}) = 1 - \alpha$$

- $\alpha \in [0, 1]$ étant le coefficient d'affaiblissement (ou de fiabilité)

Affaiblissement

Affaiblissement

Probabilité est connue avec imprécision : remplacement de la probabilité non fiable $p(x)$ par la masse $m(u)$ définie par :

$$m(u) = \alpha p(x), \quad m(\mathcal{X}) = 1 - \alpha$$

- $\alpha \in [0, 1]$ étant le coefficient d'affaiblissement (ou de fiabilité)
- Permet de prendre en compte la fiabilité relative des sources d'informations.

Affaiblissement

Affaiblissement

Probabilité est connue avec imprécision : remplacement de la probabilité non fiable $p(x)$ par la masse $m(u)$ définie par :

$$m(u) = \alpha p(x), \quad m(\mathcal{X}) = 1 - \alpha$$

- $\alpha \in [0, 1]$ étant le coefficient d'affaiblissement (ou de fiabilité)
- Permet de prendre en compte la fiabilité relative des sources d'informations.
- L'opération d'affaiblissement a pour effet de pondérer l'importance d'une source d'information lors de sa combinaison avec d'autres sources.

Affaiblissement

Affaiblissement

Probabilité est connue avec imprécision : remplacement de la probabilité non fiable $p(x)$ par la masse $m(u)$ définie par :

$$m(u) = \alpha p(x), \quad m(\mathcal{X}) = 1 - \alpha$$

- $\alpha \in [0, 1]$ étant le coefficient d'affaiblissement (ou de fiabilité)
- Permet de prendre en compte la fiabilité relative des sources d'informations.
- L'opération d'affaiblissement a pour effet de pondérer l'importance d'une source d'information lors de sa combinaison avec d'autres sources.

Objectif

Affaiblir automatiquement en chaque site $n \in \mathcal{N}$ la source d'information *a priori* (information contextuelle) lorsque celle-ci et connue avec insuffisamment de précision.

Chaîne de Markov Evidentielle Cachée à Bruit Indépendant (CMEC-BI)

Chaîne de Markov Evidentielle (CME)

Une fonction de masse m définie sur $\mathcal{P}(\mathcal{X}^N)$ est une chaîne de Markov évidentielle (CME) si elle est nulle en dehors de $[\mathcal{P}(\mathcal{X})]^N$ et si
$m(u) = m(u_1, \ldots, u_N) = m(u_1)m(u_2|u_1)\ldots m(u_N|u_{N-1})$

Chaîne de Markov Evidentielle Cachée à Bruit Indépendant (CMEC-BI)

Chaîne de Markov Evidentielle (CME)

Une fonction de masse m définie sur $\mathcal{P}(\mathcal{X}^N)$ est une chaîne de Markov évidentielle (CME) si elle est nulle en dehors de $[\mathcal{P}(\mathcal{X})]^N$ et si
$m(u) = m(u_1, \ldots, u_N) = m(u_1)m(u_2|u_1)\ldots m(u_N|u_{N-1})$

- Dans CMC-BI, la CM X est remplacée par une CME U et $q(x) \propto p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$

Chaîne de Markov Evidentielle Cachée à Bruit Indépendant (CMEC-BI)

Chaîne de Markov Evidentielle (CME)

Une fonction de masse m définie sur $\mathcal{P}(\mathcal{X}^N)$ est une chaîne de Markov évidentielle (CME) si elle est nulle en dehors de $[\mathcal{P}(\mathcal{X})]^N$ et si
$m(u) = m(u_1, \ldots, u_N) = m(u_1)m(u_2|u_1)\ldots m(u_N|u_{N-1})$

- Dans CMC-BI, la CM X est remplacée par une CME U et $q(x) \propto p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$
- La fusion s'écrit :$(m \oplus q)(x) \propto \sum_{u \in [\mathcal{P}(\mathcal{X})]^N} 1_{[x \in u]} m(u) q(x)$

Chaîne de Markov Evidentielle Cachée à Bruit Indépendant (CMEC-BI)

Chaîne de Markov Evidentielle (CME)

Une fonction de masse m définie sur $\mathcal{P}(\mathcal{X}^N)$ est une chaîne de Markov évidentielle (CME) si elle est nulle en dehors de $[\mathcal{P}(\mathcal{X})]^N$ et si
$m(u) = m(u_1, \dots, u_N) = m(u_1)m(u_2|u_1)\dots m(u_N|u_{N-1})$

- Dans CMC-BI, la CM X est remplacée par une CME U et $q(x) \propto p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$
- La fusion s'écrit :$(m \oplus q)(x) \propto \sum_{u \in [\mathcal{P}(\mathcal{X})]^N} 1_{[x \in u]} m(u) q(x)$

La fusion détruit la markovianité : impossibilité de calculer directement les marginales *a posteriori*.

Chaîne de Markov Evidentielle Cachée à Bruit Indépendant (CMEC-BI)

Chaîne de Markov Evidentielle (CME)

Une fonction de masse m définie sur $\mathcal{P}(\mathcal{X}^N)$ est une chaîne de Markov évidentielle (CME) si elle est nulle en dehors de $[\mathcal{P}(\mathcal{X})]^N$ et si
$m(u) = m(u_1, \ldots, u_N) = m(u_1)m(u_2|u_1)\ldots m(u_N|u_{N-1})$

- Dans CMC-BI, la CM X est remplacée par une CME U et $q(x) \propto p(y|x) = \prod_{n=1}^{N} p(y_n|x_n)$
- La fusion s'écrit :$(m \oplus q)(x) \propto \sum_{u \in [\mathcal{P}(\mathcal{X})]^N} 1_{[x \in u]} m(u) q(x)$

La fusion détruit la markovianité : impossibilité de calculer directement les marginales *a posteriori*.

CMEC-BI

- $(m \oplus q)(x)$ est la marginale d'une CMT particulière appelée CMEC-BI avec $p(v) \propto 1_{[x \in u]} m(u)$ et $\mathcal{U} = \mathcal{P}(\mathcal{X})$
- Les probabilités marginales *a posteriori* peuvent être calculées en considérant la CMT avant fusion.
- L'affaiblissement est estimé automatiquement par EM initialisé par k-moyennes.

Restauration d'une CM Non Stationnaire

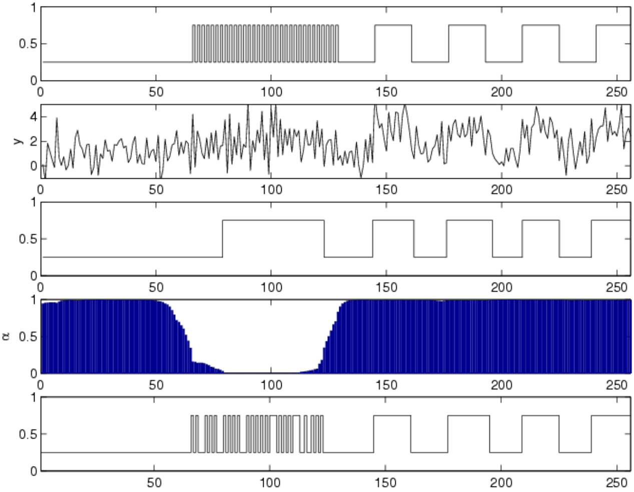


Fig.: De haut en bas : (a)processus caché, (b)processus observé, (c)résultat de la restauration non supervisée de la CMC, (d)variation de l'affaiblissement représentée par $1 - p(\mathcal{X}|y)$, (e)résultat de la segmentation non supervisée de la CMEC-BI

Segmentation d'image : comparaison CMC-BI/CMEC-BI

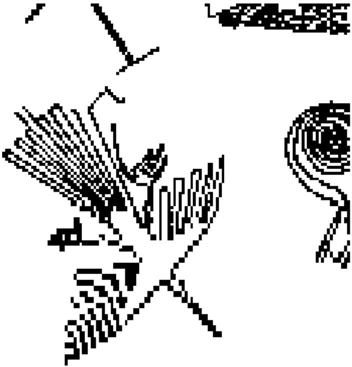


Fig.: Image originale

Segmentation d'image : comparaison CMC-BI/CMEC-BI

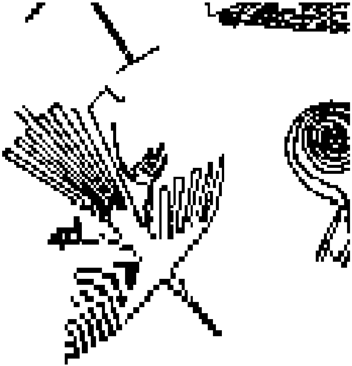


Fig.: Image originale

Fig.: Image bruitée $\mathcal{N}(0,1)$, $\mathcal{N}(2,1)$

Segmentation d'image : comparaison CMC-BI/CMEC-BI

Fig.: CMC-BI(14.6%)

Segmentation d'image : comparaison CMC-BI/CMEC-BI

Fig.: CMC-BI(14.6%)

Fig.: CMEC-BI(5.0%)

Segmentation d'image : comparaison CMC-BI/CMEC-BI

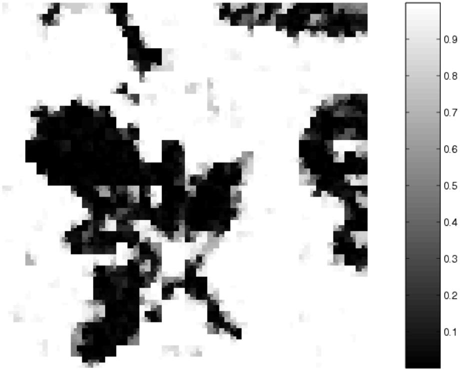


Fig.: Coefficient d'affaiblissement

Segmentation d'image : comparaison CMC-BI/CMEC-BI

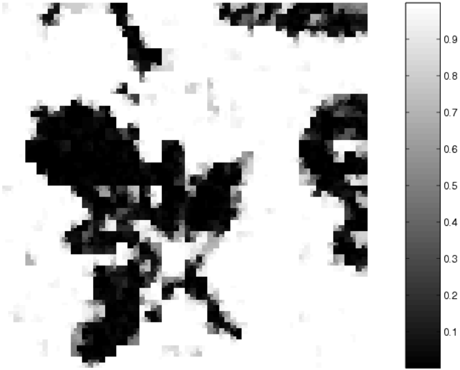


Fig.: Coefficient d'affaiblissement

- L'affaiblissement est plus fort pour les zones comprenant des détails fins

Comparaison avec le modèle CMM-NSC-BI

CMM-NSC-BI

- On considère différentes zones de stationnarités correspondant chacune à un état auxiliaire.

Comparaison avec le modèle CMM-NSC-BI

CMM-NSC-BI

- On considère différentes zones de stationnarités correspondant chacune à un état auxiliaire.

CMEC-BI

- L'influence de la loi *a priori* et donc de l'information contextuelle est affaiblie par rapport à l'information issue des observations dans les zones présentant des détails fins.
- Le modèle CMEC-BI permet donc d'adapter automatiquement l'influence de l'information contextuelle en fonction des caractéristiques de l'image.

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions :

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions :

Multispectrales

Différentes longueurs d'ondes :

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions :

Multispectrales

Différentes longueurs d'ondes :

Structures diffuses

Nuages, poussières :

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions : Arbre de Markov caché

Multispectrales

Différentes longueurs d'ondes :

Structures diffuses

Nuages, poussières :

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions : Arbre de Markov caché

Multispectrales

Différentes longueurs d'ondes : Loi d'observation multivariée

Structures diffuses

Nuages, poussières :

Caractéristique des images astronomiques

Multirésolutions

Images acquisent à différentes résolutions : Arbre de Markov caché

Multispectrales

Différentes longueurs d'ondes : Loi d'observation multivariée

Structures diffuses

Nuages, poussières : Modélisation floue

Approche floue

Pour $\mathcal{X} = \{0, 1\}$, existence de sites contenant un mélange des deux classes sans qu'aucune des deux classes ne s'impose.

- $x_s = 0$ si s appartient à la classe "0"
- $x_s = 1$ si s appartient à la classe "1"
- $x_s = \varepsilon_s \in]0, 1[$ si la proportion de la classe "0" dans le site est $\varepsilon_s \in]0, 1[$.

On obtient une appartenance 'floue" du site

Approche floue

Pour $\mathcal{X} = \{0, 1\}$, existence de sites contenant un mélange des deux classes sans qu'aucune des deux classes ne s'impose.

- $x_s = 0$ si s appartient à la classe "0"
- $x_s = 1$ si s appartient à la classe "1"
- $x_s = \varepsilon_s \in]0, 1[$ si la proportion de la classe "0" dans le site est $\varepsilon_s \in]0, 1[$.

On obtient une appartenance 'floue" du site

Loi de X_s

Donnée par une densité $p(x_s)$ par rapport à une mesure ν incluant des composantes discrètes et continues

$$\nu = \delta_0 + \delta_1 + \mu$$

et

$$p(0) + p(1) + \int_0^1 p(\varepsilon)d\varepsilon = 1$$

Arbre de Markov flou

Notations

$\mathcal{S}$: ensemble fini d'indices
$\mathcal{S}^1, \ldots, \mathcal{S}^N$: partition de $\mathcal{S}$
$\mathcal{S}^1$: racine de l'arbre
s^- : parent d'un site s
s^+ : enfants de s
s^{++} : ensemble des descendants de s

Arbre de Markov flou

Notations

$\mathcal{S}$: ensemble fini d'indices
$\mathcal{S}^1, \ldots, \mathcal{S}^N$: partition de $\mathcal{S}$
$\mathcal{S}^1$: racine de l'arbre
s^- : parent d'un site s
s^+ : enfants de s
s^{++} : ensemble des descendants de s

- Nous considérons à présent un arbre de Markov $X = (X)_{s \in \mathcal{S}}$ avec $X_s \in [0, 1]$

Arbre de Markov flou

Notations

$\mathcal{S}$: ensemble fini d'indices
$\mathcal{S}^1, \dots, \mathcal{S}^N$: partition de $\mathcal{S}$
$\mathcal{S}^1$: racine de l'arbre
s^- : parent d'un site s
s^+ : enfants de s
s^{++} : ensemble des descendants de s

- Nous considérons à présent un arbre de Markov $X = (X)_{s \in \mathcal{S}}$ avec $X_s \in [0, 1]$
- La loi d'un arbre de Markov flou X par rapport à la mesure ν^S est définie par

$$p(x) = p(x_1) \prod_{n=2}^{N} \prod_{s \in \mathcal{S}^n} p(x_s | x_{s^-})$$

Arbre de Markov flou multispectral

- La distribution de $Y_s = (Y_s^{(1)}, Y_s^{(2)}, \ldots, Y_s^{(D)})$ conditionnellement à x_s est modélisée par une densité gaussienne multivariée $p(y_s|x_s)$

Arbre de Markov flou multispectral

- La distribution de $Y_s = (Y_s^{(1)}, Y_s^{(2)}, \dots, Y_s^{(D)})$ conditionnellement à x_s est modélisée par une densité gaussienne multivariée $p(y_s|x_s)$
- Le vecteur moyenne μ_{ε_s} et la matrice Γ_{ε_s} associées au site flou $X_s = \varepsilon_s$ dépendent linéairement des paramètres durs (μ_0, μ_1), (Γ_0, Γ_1) :

$$\left\{ \begin{array}{l} \mu_{\varepsilon_s} = (1-\varepsilon_s)\mu_0 + \varepsilon_s \mu_1 \\ \Gamma_{\varepsilon_s} = (1-\varepsilon_s)^2 \Gamma_0 + \varepsilon_s^2 \Gamma_1 \end{array} \right.$$

Arbre de Markov flou multispectral

- La distribution de $Y_s = (Y_s^{(1)}, Y_s^{(2)}, \dots, Y_s^{(D)})$ conditionnellement à x_s est modélisée par une densité gaussienne multivariée $p(y_s|x_s)$
- Le vecteur moyenne μ_{ε_s} et la matrice Γ_{ε_s} associées au site flou $X_s = \varepsilon_s$ dépendent linéairement des paramètres durs (μ_0, μ_1), (Γ_0, Γ_1) :

$$\left\{ \begin{array}{ll} \mu_{\varepsilon_s} & = (1-\varepsilon_s)\mu_0 + \varepsilon_s \mu_1 \\ \Gamma_{\varepsilon_s} & = (1-\varepsilon_s)^2 \Gamma_0 + \varepsilon_s^2 \Gamma_1 \end{array} \right.$$

Loi de Z

$$p(z) = p(x_1)p(y_1|x_1) \prod_{n=2}^{N} \prod_{s \in S^n} p(x_s|x_{s^-})p(y_s|x_s)$$

Calcul des marginales a posteriori

Passe montante : pour $n = N \ldots 2$, $\forall s \in \mathcal{S}^n$

$$\beta_{s,s^-}(z_{s^-}) \propto \begin{cases} \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-})p(y_s|x_s)d\nu(x_s) & \text{si } n = N \\ \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-})p(y_s|x_s)\prod_{t \in s^+} \beta_{t,s}(z_s)d\nu(x_s) & \text{sinon} \end{cases}$$

$$p(x_s|x_{s^-}, y) \propto \begin{cases} p(x_s|x_{s^-})p(y_s|x_s) & \text{si } n = N \\ p(x_s|x_{s^-})p(y_s|x_s)\prod_{t \in s^+} \beta_{t,s}(z_s) & \text{sinon} \end{cases}$$

Calcul des marginales a posteriori

Passe montante : pour $n = N \dots 2$, $\forall s \in \mathcal{S}^n$

$$\beta_{s,s^-}(z_{s^-}) \propto \begin{cases} \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-})p(y_s|x_s)\,d\nu(x_s) & \text{si } n = N \\ \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-})p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s)\,d\nu(x_s) & \text{sinon} \end{cases}$$

$$p(x_s|x_{s^-}, y) \propto \begin{cases} p(x_s|x_{s^-})p(y_s|x_s) & \text{si } n = N \\ p(x_s|x_{s^-})p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s) & \text{sinon} \end{cases}$$

Racine

$$p(x_1|y) \propto p(z_1) \prod_{t \in 1^+} \beta_{t,1}(x_1)$$

Calcul des marginales a posteriori

Passe montante : pour $n = N \dots 2$, $\forall s \in \mathcal{S}^n$

$$\beta_{s,s^-}(z_{s^-}) \propto \begin{cases} \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-}) p(y_s|x_s) d\nu(x_s) & \text{si } n = N \\ \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-}) p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s) d\nu(x_s) & \text{sinon} \end{cases}$$

$$p(x_s|x_{s^-}, y) \propto \begin{cases} p(x_s|x_{s^-}) p(y_s|x_s) & \text{si } n = N \\ p(x_s|x_{s^-}) p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s) & \text{sinon} \end{cases}$$

Racine

$$p(x_1|y) \propto p(z_1) \prod_{t \in 1^+} \beta_{t,1}(x_1)$$

pour $n = 2 \dots N$, $\forall s \in \mathcal{S}^n$

$$p(x_s, x_{s^-}|y) = p(x_{s^-}|y) p(x_s|x_{s^-}, y)$$
$$p(x_s|y) = \int_{x_{s^-} \in \mathcal{X}} p(x_s, x_{s^-}|y) d\nu(x_{s^-})$$

Calcul des marginales a posteriori

Passe montante : pour $n = N \dots 2$, $\forall s \in \mathcal{S}^n$

$$\beta_{s,s^-}(z_{s^-}) \propto \begin{cases} \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-}) p(y_s|x_s) d\nu(x_s) & \text{si } n = N \\ \int_{x_s \in \mathcal{X}} p(x_s|x_{s^-}) p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s) d\nu(x_s) & \text{sinon} \end{cases}$$

$$p(x_s|x_{s^-}, y) \propto \begin{cases} p(x_s|x_{s^-}) p(y_s|x_s) & \text{si } n = N \\ p(x_s|x_{s^-}) p(y_s|x_s) \prod_{t \in s^+} \beta_{t,s}(z_s) & \text{sinon} \end{cases}$$

Racine

$$p(x_1|y) \propto p(z_1) \prod_{t \in 1^+} \beta_{t,1}(x_1)$$

pour $n = 2 \dots N$, $\forall s \in \mathcal{S}^n$

$$p(x_s, x_{s^-}|y) = p(x_{s^-}|y) p(x_s|x_{s^-}, y)$$
$$p(x_s|y) = \int_{x_{s^-} \in \mathcal{X}} p(x_s, x_{s^-}|y) d\nu(x_{s^-})$$

- Approximation numérique : Intervalle $[0, 1]$ divisé en F sous-intervalles flous $[a_0 = 0, a_1 = \frac{1}{F}], [a_1 = \frac{1}{F}, a_2 = \frac{2}{F}], \dots, [a_{F-1} = \frac{F-1}{F}, a_F = 1]$.

Image de synthèse

- Image de synthèse trispectrale. La“vérité-terrain” est un champ de Markov flou à 10 niveaux. le taux de flou = 53.96%

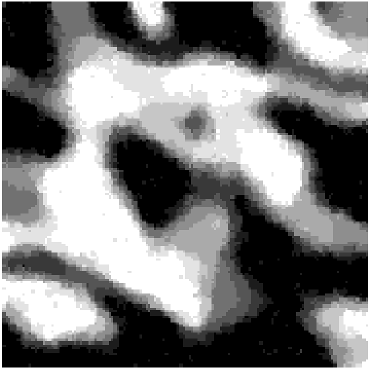
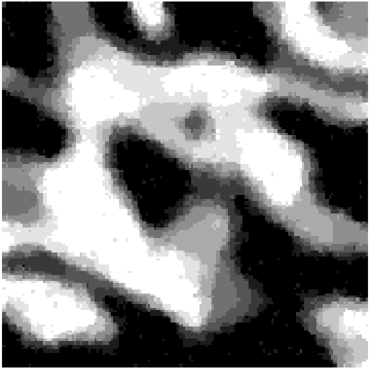


Image de synthèse

- Image de synthèse trispectrale. La"vérité-terrain" est un champ de Markov flou à 10 niveaux. le taux de flou = 53.96%



- Les paramètres sont estimés par un algorithme EM initialisé par k-moyennes.

Résultat de la segmentation

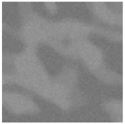


Fig.: 1ere bande

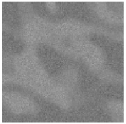


Fig.: 2e bande

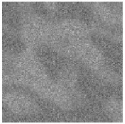


Fig.: 3e bande

Résultat de la segmentation

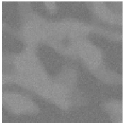


Fig.: 1ere bande

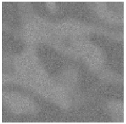


Fig.: 2e bande

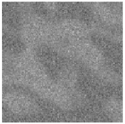


Fig.: 3e bande

Fig.: 8 niveaux de flou et 2 classes dures, 4,58% d'erreur, 44.96% de flou

Résultat de la segmentation

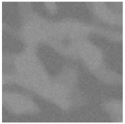


Fig.: 1ere bande

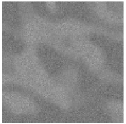


Fig.: 2e bande

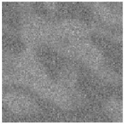


Fig.: 3e bande

Fig.: 8 niveaux de flou et 2 classes dures, 4,58% d'erreur, 44.96% de flou

- L'algorithme restitue correctement l'image d'origine avec un taux d'erreur de 4.58%

Résultat de la segmentation

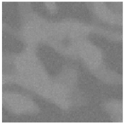


Fig.: 1ere bande

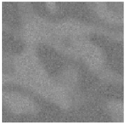


Fig.: 2e bande

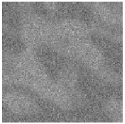


Fig.: 3e bande

Fig.: 8 niveaux de flou et 2 classes dures, 4,58% d'erreur, 44.96% de flou

- L'algorithme restitue correctement l'image d'origine avec un taux d'erreur de 4.58%
- L'information floue est correctement restituée 44.96% vs 53.96%

Image astronomique

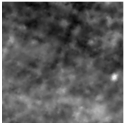


Fig.: 60μm

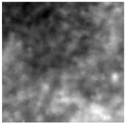


Fig.: 100μm

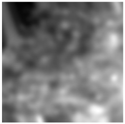


Fig.: 170μm

- Objectif : mettre en évidence la structure filamentaire de poussière interstellaire gênant l'observation de champs de haute altitude de notre galaxie

Image astronomique

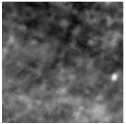


Fig.: 60μm

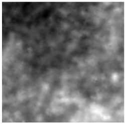


Fig.: 100μm

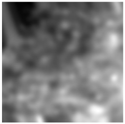


Fig.: 170μm

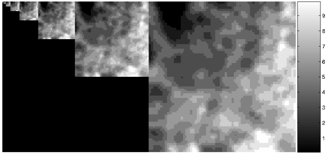


Fig.: 8 niveaux de flou et 2 classes dures

- Objectif : mettre en évidence la structure filamentaire de poussière interstellaire gênant l'observation de champs de haute altitude de notre galaxie
- Ne permet pas de mettre en évidence la structure filamentaire de la poussière interstellaire
- Restitue une gradation floue entre les régions homogènes

Conclusions et perspectives

Contributions

Conclusions et perspectives

Contributions

- Algorithme d'estimation EM dans le cas des chaînes de Markov Couples.

Conclusions et perspectives

Contributions

- Algorithme d'estimation EM dans le cas des chaînes de Markov Couples.
- Etude d'un cas particulier de chaîne couple partiellement de Markov permettant la segmentation de processus gaussiens à corrélation longue.

Conclusions et perspectives

Contributions

- Algorithme d'estimation EM dans le cas des chaînes de Markov Couples.
- Etude d'un cas particulier de chaîne couple partiellement de Markov permettant la segmentation de processus gaussiens à corrélation longue.
- Proposition d'un cas particulier de chaîne triplet permettant la segmentation non supervisée de processus non stationnaires.

Conclusions et perspectives

Contributions

- Algorithme d'estimation EM dans le cas des chaînes de Markov Couples.
- Etude d'un cas particulier de chaîne couple partiellement de Markov permettant la segmentation de processus gaussiens à corrélation longue.
- Proposition d'un cas particulier de chaîne triplet permettant la segmentation non supervisée de processus non stationnaires.
- Etude des possibilités d'extension des modèles probabilistes classiques à un modèle « évidentiel », avec la loi *a posteriori* du processus caché donnée par la fusion de Dempster-Shafer.

Conclusions et perspectives

Contributions

- Algorithme d'estimation EM dans le cas des chaînes de Markov Couples.
- Etude d'un cas particulier de chaîne couple partiellement de Markov permettant la segmentation de processus gaussiens à corrélation longue.
- Proposition d'un cas particulier de chaîne triplet permettant la segmentation non supervisée de processus non stationnaires.
- Etude des possibilités d'extension des modèles probabilistes classiques à un modèle « évidentiel », avec la loi *a posteriori* du processus caché donnée par la fusion de Dempster-Shafer.
- Extension des modèles de chaînes de Markov floues et de champs de Markov flous à un modèle d'arbres de Markov flou caché multispectraux

Conclusions et perspectives

Perspectives

- Extension des CCPM gaussiennes au cas non centré.

Conclusions et perspectives

Perspectives

- Extension des CCPM gaussiennes au cas non centré.
- Extension du modèle d'arbres de Markov cachés flous multispectraux à un modèle d'arbres de Markov couples flous multispectraux

Conclusions et perspectives

Perspectives

- Extension des CCPM gaussiennes au cas non centré.
- Extension du modèle d'arbres de Markov cachés flous multispectraux à un modèle d'arbres de Markov couples flous multispectraux
- Etudes des liens entre les modèles de Markov conditionnels et les modèles triplet partiellement de Markov.